



QY 781 gggaactattagctatgcccattgctgaattgctgtttcacagaggattgatagagaata 840  
 Db 811 GGGAACTATTAGCTATGCCATTGCTGAATTGGTTTCACAGGATTGATAGTGAAGATA 870  
 QY 841 aattagacaccttcaatataccacagctatttgcataccttgaggaagtgaaagatatag 900  
 Db 871 AATTAGACACCTTCAATATACCCAGCTATTTTGGCATCTACTTGAGGAAGTGAAAGATATAG 930  
 QY 901 tggagaggagcagatcattcacaaatgatcatatagaggggtttgatcttgatagcgtag 960  
 Db 931 TGGAGAGGACCGATCATTACAAATGATCATATAGAGGGGTTTGATCTTGATAGCGGTAG 990  
 QY 961 aaatgcagaggaatgataaaatgggttagaggggaaagtgttaccaggttgcaggccct 1020  
 Db 991 AAATGCAGGAGAAATGATAAATGGGTAGAGGGGAAAGTTTACCAAGGTTGTCAGGGCCT 1050  
 QY 1021 tcacagagcctataatttcaaacacagtttgacactgaaatcatggaacaaactatatgaca 1080  
 Db 1051 TCACAGAGCCTATAATTTCAACACAGTTTGGACCTGGAATCATGGCAAACTATATGACA 1110  
 QY 1081 aattcactcacatgttagttttagaattggaggaagaaagctacggagagacacaaagtatca 1140  
 Db 1111 AATTCACTCACATTGTAGTTTTCAGATTGGAGCAAGAACTACCGAAGACACCAAGTATCA 1170  
 QY 1141 tctagtgcttcccaagattgatgatagtttttagttgtgtggaataaactgtgtgcc 1200  
 Db 1171 TCCTAGTGCTTTCCCAAGATTGATGATAGTTTGTAGTTGTTGGAATAAAGTGTGCTC 1230  
 QY 1201 ctatcacatatgcactagaggttgcgaatgattgtcacagaagaagatttgagg 1260  
 Db 1231 CTATCACATATATGCTATGAGGGTTTTCGAATGATTTCAAGAAATTTGAGAGG 1290  
 QY 1261 ggtcaaatatgaaagcatttgccttgtgtgagagagaatgtttcttctgatttaaat 1320  
 Db 1291 GGTCAAAATATCAAGCAATTGTGCTCTGTGGAGAGAGAAATGTTTCTTGATTAAAT 1350  
 QY 1321 ctgdatcccaaatcgttaattgtgggaagaataatgaaagltgacatgaaattttaa 1380  
 Db 1351 CNGTGATACCCAAATCGTAATGTGGGAAGAAATGAGAAGTTGAACATGAAATTTTAA 1410  
 QY 1381 aaaaaaataaaaaaataaaaaa 1408  
 Db 1411 AAAAAAATAAAAAAATAAAAAA 1438  
 RESULT 2  
 LOCUS AB049752 1437 bp mRNA PLN 16-MAR-2001  
 DEFINITION Atropa belladonna AbsAMT1 mRNA for S-adenosyl-L-methionine:salicylic acid carboxyl methyltransferase, complete cds.  
 ACCESSION AB049752  
 VERSION AB049752.1 GI:13366160  
 KEYWORDS Atropa belladonna (strain:M8) root cDNA to mRNA, clone:lambda AbsAMT1.  
 ORGANISM Atropa belladonna  
 Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I; Solanales; Solanaceae; Atropa.  
 REFERENCE 1 (sites)  
 AUTHORS Fukami,H., Asakura,T., Hirano,H., Abe,K., Shimomura,K. and Yamakawa,T.  
 TITLE Cloning and expression of salicylic acid inducible and active S-adenosyl-L-methionine:salicylic acid carboxyl methyltransferase in transformed root culture of Atropa belladonna  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 1437)  
 AUTHORS Fukami,H., Asakura,T., Hirano,H., Abe,K., Shimomura,K. and Yamakawa,T.  
 TITLE Direct Submission  
 JOURNAL Submitted (10-OCT-2000) Takashi Yamakawa, The University of Tokyo,

Department of Global Agricultural Sciences, 1-1, Yayoi 1-chome, Bunkyo-ku, Tokyo 113-8657, Japan  
 (E-mail:ayama@mail.ecc.u-tokyo.ac.jp, Tel.81-3-5841-7515, Fax:81-3-5841-5304)  
 FEATURES  
 source Location/Qualifiers  
 1. .1437  
 /organism="Atropa belladonna"  
 /strain="M8"  
 /db\_xref="taxon:33113"  
 /clone="lambda AbsAMT1"  
 /tissue\_type="root"  
 62. .1135  
 /gene="AbsAMT1"  
 62. .1135  
 /gene="AbsAMT1"  
 /codon\_start=1  
 /product="S-adenosyl-L-methionine:salicylic acid carboxyl methyltransferase"  
 /protein\_id="BAB39396.1"  
 /db\_xref="GI:13366161"  
 /translation="MKVEVLMHNGGNDISYANNSLVORVKVILMTKPITEOAISDLV CSFPEPLCIADLQCSGGANTFLVSELVIVKEKRIHNLQAGNLFHNDLPNDP NTIPOSIGKFOQDLRKQIGEEFGCPSPGVSFVTRLPSPSLHVSYSIMWLISO VPDLEAKKENIYIASTSPSPVIAKAYKQYKDFSNFLAYRSEELMKGGKMWLTFLGR ESDPSSKECCYIWEILLMALNELVLEELIEEKYDSFNIPQITPSPVKKVIVKEG SFINRLEATRVHNNVSNNEGNGYVAKOMRAVEPLLVSOPDQKLMNLVFKYVEEI ISDCISKEKTEFINVIVSLTKR"  
 BASE COUNT 488 a 225 c 290 g 434 t  
 ORIGIN  
 Query Match 16.7%; Score 238.2; Bb 12; Length 1437;  
 Best Local Similarity 54.9%; Pred. No. 3; Be-41;  
 Matches 599; Conservative 0; Mismatches 453; Indels 39; Gaps 5;  
 QY 80 aaagtgaaacaaagtgtgttcatacaacagggggagaggaagagatttatcacaaaa 139  
 Db 64 GAAGGTTTGTGAAGTTCTTCACATGAATGAGAGAAATGGTGACATTAGCTATGCAACAA 123  
 QY 140 ctctctcttcacgaacaaagtggcctcaatggcacagccagcgtacgaataatgcagtga 199  
 Db 124 TTCTCTGGTTCAGAGAAAGTGTAATTCATGACAAAGCCATAACAGAACAAAGCCATAAG 183  
 QY 200 aactctcttcacgaagatttccaccttcacagctcttcaacgacgagcacttgggttgc 259  
 Db 184 TGATCTCTACTGCAGC---TTCTTCCAGAGAAACCTTATGCAATTCCTGATTGGGTGTC 240  
 QY 260 aaggggtccaaacacattgcagtgattttctacgatcaagagaatgatgaaagaatg 319  
 Db 241 TTCTGAGAGGAAACATTTTGTGATGATGAGAACTGTGTAATAATCGTCAAAAAGAAAG 300  
 QY 320 cagggdaattgaattgcacaaacacttgcaggttcaattgaatgatctttttggaaa 379  
 Db 301 AAAATAATATATCTTCATCTCACTCAGCGGGAATCTTTTTCACITCAATGATCTCTCTGSCAA 360  
 QY 380 tgatttcataacctcttccaa-----agccctgcgtctgaggttatggtaacaaat 432  
 Db 361 TGATTTTAAACACCATTTTTCAGTCATGGGGAATTTTCAACAGAGATTGAGAAAGCAAT 420  
 QY 433 gtagaga-----agttcgtgtatgtgatggagatcccggtgtcttccatggcggt 487  
 Db 421 TGAGAGAAATTTGTGTCATGTTTITTAGTGGAGTGCTGCTGCTCATTTTATATATACTA 480  
 QY 488 ttttctcgtacacagcttaccattagttcattctcttaccagtttccattggttactca 547  
 Db 481 TTTCCTCTCAGAGAGTTTGCATTTGTGTCACTCCAGTTAAGTCTATGTSCTATATCA 540  
 QY 548 ggcacacaaagagctcacaaacagcagaagcgttggcattaaacaggggagagattacat 607  
 Db 541 AGTCTCTGATTAAATTGAAAG-----AACCAAGGAGCAACATTTACAT 582  
 QY 608 atcaagacaaagcctctctgtgtaagagaagccttacttctcaatttcatgaagatt 667

GenCore version 4.5  
Copyright (c) 1994 - 2000 Compugen Ltd

CM protein - protein search, using sw model

Run on: August 31, 2001, 15:00:32 ; Search time 59.12 seconds

(without alignments)  
796,895 Million cell updates/sec

Title: US-09-577-657A-1  
Perfect score: 1847  
Sequence: 1 FMNRGESSYAONSSFTQ2 LEAKIPKTTIILVLKICG 356

Scoring table:  
BLOSUM62  
Gapop 10 0, Gapext 0 5

Searched: 425026 seqs, 12395027 residues  
Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database:

SPRENBL\_16:  
1: sp\_archaea:  
2: sp\_bacteria:  
3: sp\_fungi:  
4: sp\_human:  
5: sp\_invertebrate:  
6: sp\_mammal:  
7: sp\_mhc:  
8: sp\_organelle:  
9: sp\_phage:  
10: sp\_plant:  
11: sp\_procent:  
12: sp\_unclassified:  
13: sp\_vertebrate:  
14: sp\_virus:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1847	100.0	369	10 Q9FZN8	Q9FZN8 camellia sl
2	724	39.2	392	10 Q9SBK6	Q9SBK6 brassica ra
3	713	38.6	359	10 Q9SPV4	Q9SPV4 clarkia dre
4	637.5	34.5	354	10 Q9FJ26	Q9FJ26 arabidopsis
5	630	34.1	371	10 Q9J234	Q9J234 arabidopsis
6	621.5	33.6	364	10 Q9FT29	Q9FT29 arabidopsis
7	585.5	31.7	363	10 Q9IS20	Q9IS20 arabidopsis
8	582.5	31.5	368	10 Q9IPR5	Q9IPR5 arabidopsis
9	570	30.9	359	10 Q9ZPT3	Q9ZPT3 arabidopsis
10	499.5	27.0	351	10 Q9XIS7	Q9XIS7 arabidopsis
11	496.5	26.9	348	10 Q9FYC4	Q9FYC4 arabidopsis
12	485.5	26.3	348	10 Q9FYC6	Q9FYC6 arabidopsis
13	482.5	26.1	359	10 Q9IS10	Q9IS10 arabidopsis
14	482	26.0	374	10 Q9FKC8	Q9FKC8 arabidopsis
15	481	26.0	386	10 Q9ELN8	Q9ELN8 arabidopsis
16	480	26.0	379	10 Q9FYC3	Q9FYC3 arabidopsis
17	480	25.0	319	10 Q9F552	Q9F552 arabidopsis
18	475.5	25.7	323	10 Q9FWK4	Q9FWK4 arabidopsis
19	475.5	25.7	361	10 Q9FKK0	Q9FKK0 arabidopsis

20	466.5	25.3	380	10 Q9FWJ1	Q9FWJ1 oryza sativ
21	454	24.6	362	10 Q9FKU0	Q9FKU0 arabidopsis
22	436.5	23.6	380	10 Q9FWA2	Q9FWA2 arabidopsis
23	424	23.0	353	10 Q9MAF2	Q9MAF2 brassica na
24	387	21.0	318	10 Q9FT13	Q9FT13 brassica na
25	105.5	5.7	100	10 Q9LS11	Q9LS11 arabidopsis
26	101.5	5.5	1016	5 Q9UT45	Q9UT45 schizosacch
27	99	5.4	319	5 Q9Z637	Q9Z637 caenorhabdi
28	98	5.3	1236	5 Q44875	Q44875 caenorhabdi
29	97	5.3	1694	5 Q21436	Q21436 caenorhabdi
30	96	5.2	1280	2 Q9ZAA5	Q9ZAA5 clostridium
31	96	5.2	1862	10 Q9SHM5	Q9SHM5 arabidopsis
32	94.5	5.1	469	14 Q9WBS2	Q9WBS2 physalis se
33	94.5	5.1	590	2 Q54535	Q54535 streptococ
34	94.5	5.1	787	1 Q58025	Q58025 pyrococcus
35	94.5	5.1	3125	14 Q9IFR9	Q9IFR9 turkey herp
36	94.5	5.1	3442	14 Q5E6N3	Q5E6N3 turkey herp
37	94	5.1	983	10 Q9ZUK3	Q9ZUK3 arabidopsis
38	93.5	5.1	525	1 Q5X840	Q5X840 schizosacch
39	93	5.0	295	2 Q9XDK6	Q9XDK6 bacteroides
40	93	5.0	736	5 Q25856	Q25856 plasmodium
41	93	5.0	749	5 Q02602	Q02602 plasmodium
42	93	5.0	897	10 Q9LHC0	Q9LHC0 arabidopsis
43	92.5	5.0	2185	14 Q9E7C3	Q9E7C3 human coxa
44	92.5	5.0	2185	14 Q9E7C2	Q9E7C2 human coxa
45	92	5.0	392	3 Q94240	Q94240 schizosacch

## ALIGNMENTS

RESULT	1	PRELIMINARY:	PRT:	369 AA.
Q9FZN8	Q9FZN8			
ID	Q9FZN8			
AC	Q9FZN8			
DT	01-MAR-2001 (TREMBlrel. 16, Created)			
DT	01-MAR-2001 (TREMBlrel. 16, Last sequence update)			
DT	01-MAR-2001 (TREMBlrel. 16, Last annotation update)			
DE	CAFFEINE SYNTHASE.			
GN	TCS1			
OS	Camellia sinensis (Tea).			
OC	Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;			
OC	Magnoliophyta; eudicotyledons; core eudicots; Asteridae; Ericales;			
OC	Theaceae; Camellia.			
OX	NCBI_TaxID=4442;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=2043735; PubMed=10984041;			
PA	Kato M, Mizuno K, Crozier A, Fujimura T, Ashihara H.;			
RT	"Caffeine synthase gene from tea leaves.";			
RL	Nature 406:956-957(2000).			
DR	EMBL: AB031280; BMB12278.1;			
SQ	SEQUENCE 369 AA; 41772 MW; E6D26J087E475576 CRC64;			
Query Match	100.0%; Score 1847; DB 10; Length 369;			
Best Local Similarity	100.0%; Pred. No. 1.1e-144;			
Matches 350; Conservative	U, Mismatches U, Indels 0, Gaps 0;			
QY	1 FMNRGESSYAONSSFTQVASMALPNAVEFLFSRDFHLQALNADGCAAGPNTF 60			
DB	14 FMNRGESSYAONSSFTQVASMALPNAVEFLFSRDFHLQALNADGCAAGPNTF 73			
QY	51 AVISTIKRMMEKKCPRLNCGTLEIYVINDLFGNPNFLFKGLSSEVIGNCEEPVCM 120			
DB	74 AVISTIKRMMEKKCPRLNCGTLEIYVINDLFGNPNFLFKGLSSEVIGNCEEPVCM 133			
QY	121 GVSSEFGRGLFPFNSLHLVSSYSVHMLTQAPKGLTSEGLALNKGLTIKTSPPVPE 180			
DB	134 GVSSEFGRGLFPFNSLHLVSSYSVHMLTQAPKGLTSEGLALNKGLTIKTSPPVPE 193			
QY	181 AVTSQHFRTMTLAPSSFVVPNNGMTLTPPGCTSPSCMGCSFTWELLAMATAETVS 240			

Db 194 AYLSOFHEDFTMFENARSOEVVPCNGLLRGRCSPSPDSOCFTWELLAMATAELVS 253  
 QY 241 OGLIDEDKLDTPNIPSYASLEEVKDIYERDGSFTIDHIEGFDLDSVEMQENDKWKVPGCK 300  
 Db 254 OGLIDEDKLDTPNIPSYASLEEVKDIYERDGSFTIDHIEGFDLDSVEMQENDKWKVPGCK 313  
 QY 301 FTKVRAATEPTISNQGPELMDKLYDKFTHVVSOLFAPKPKTTSILVLSKIDG 356  
 Db 314 FTKVRAATEPTISNQGPELMDKLYDKFTHVVSOLFAPKPKTTSILVLSKIDG 369

RESULT 2  
 Q9SRK6 PRELIMINARY: PRT: 392 AA.  
 ID O9SRK6  
 AC O9SRK6  
 DT 01-MAY-2000 (TREMBlrel. 13, Created)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
 DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)  
 DE FLORAL NECTARY-SPECIFIC PROTEIN.  
 GN NTR1.  
 OS Brassica rapa subsp. pekinensis (Chinese cabbage) (Celery cabbage).  
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;  
 OC Brassicales; Brassicaceae; Brassica.  
 OX NCBI\_TaxID=51351;  
 RN 11  
 RP SEQUENCE FROM N.A.  
 RA Song J.T., Seo H.S., Song S.L., Lee J.S., Choi Y.D.;  
 RT "Characterization of a novel gene expressed specifically in the floral  
 nectaries of Brassica campestris L. ssp. pekinensis.";  
 RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.  
 DF EMBL: AF179222; AAF22389.1; ...  
 SQ SEQUENCE 392 AA, 43815 MW, 253785305365757 CRC64;

Query Match 39.2%; Score 724; DB 10; Length 392;  
 Best local Similarity 38.1%; Pred. No. 7.9e-52;  
 Matches 146; Conservative 78; Mismatches 129; Indels 30; Gaps 5;

QY 2 MNRGESSYAONSFTQOVASMAQPALENAVETLFSKDHLOALNADLGCAGPNTFA 61  
 Db 9 MNKGNGETSTAKNSIVOSNITSIGRRVMDALKLMIRNSELISFGIADLGSSGPNLSL 68  
 QY 62 VISTIKRMEKKCKELNCUTLELVYVNLDFGNDFTLFGKLS-----EVGNKCEV- 115  
 Db 69 SISNIVETIONLCCHDRPVPBELSLNDLPNDENITFASLPEFYDRVKKRNNNESLG 128  
 QY 116 -----PCYVMGVPGSFHRLPRNSLHLVHSSYSVHMLTQAPK- LISRGLAL---- 163  
 Db 129 FEHOGSGPCFVSAPGSFYGRLEPRRSILHFVHSSSLHMLSOVPGCEVNNKDGCVITADL 188  
 QY 164 -NNGKIYISKTSPPVYVEALYSOFHEDFTMFENARSOEVVPCNGLLRGRCSPSPDSM 222  
 Db 189 DNGKITYLSTKSPKSAKNAKVALOPQTFDFVFLSPRSEFELVPGSRMVLSTLGRSSPDPTTE 248  
 QY 223 OSCFTWELLAMATAELVSOGCLIDEDKLDTPNIPSYASLEEVKDIYERDGSFTIDHIEGF 282  
 Db 249 ESCCYOWELLAOALMSLAKESGIIIEENIDAFNADYAAAPPELKMALIEKGSFSDLELS 308  
 QY 283 DLD-----SVEMQENDKWKVGEKFTKYVRAATEPTISNQGPELMDKLYDKFTH 331  
 Db 309 PVMGEGSISDSDSYDVRKPRPEALASGRVAKTIRAVVEPMLEPTFGQVMDLEFERVAK 368  
 QY 332 IVVSOLFAPKPKTTSILVLSKIDG 354  
 Db 369 LVGEYVVSPPRTIYIVLSLRM 391

RESULT 3  
 Q9SPV4 PRELIMINARY: PRT: 359 AA.  
 ID O9SPV4  
 AC O9SPV4  
 DT 01-MAY-2000 (TREMBlrel. 13, Created)

DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)  
 DE S-ADENOSYL-L-METHIONINE:SALICYLIC ACID CARBOXYL METHYLTRANSFERASE.  
 GN SAMT.  
 OS Clarkia breweri.  
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;  
 OC Myrtales; Onagraceae; Clarkia.  
 OX NCBI\_TaxID=36903;  
 RN 11  
 RP SEQUENCE FROM N.A.  
 RA Ross J.P., Nam K.H., D'Auria J., Pichersky E.;  
 RT "S-adenosyl-L-methionine:salicylic acid carboxyl methyltransferase, an  
 enzyme involved in floral scent production and plant defense,  
 represents a new class of plant methyltransferases.";  
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.  
 DF EMBL: AF133053; AAF00108.1; ...  
 KM Transferase; Methyltransferase.  
 SQ SEQUENCE 359 AA, 40289 MW, BOFE3E41AE8EDB51 CRC64;

Query Match 38.6%; Score 713; DB 10; Length 359;  
 Best local Similarity 41.5%; Pred. No. 5.7e-51;  
 Matches 152; Conservative 68; Mismatches 110; Indels 36; Gaps 8;

QY 2 MNRGESSYAONSFTQOVASMAQPALENAVETLFSKDHLOALNADLGCAGPNT-F 60  
 Db 9 MKGAGENSYVAMNSFIQROVISTITKPTTEAATLALSGTIVTRIALADLGSSGNALF 68  
 QY 64 AV---ISTIKRMEKKCKELNCUTLELVYVNLDFGNDFTLFGKLSSEVGNKCEVPC 117  
 Db 69 ATELTKTYEELRKMKGR---NSPEYQIFLNDKPGNFALFISLPTF---NDVPGV-C 121  
 QY 118 YMGVPGSFHRLPRNSLHLVHSSYSVHMLTQAPKGLTSRGLALNKGKIYSTSPV 177  
 Db 122 FINGVGSFYGLLPRNTLHFIHSSYSMLMSQVIGIES-----NKGNTIMANTQPOS 175  
 QY 178 VREAYISOFHEDFTMFENARSOEVVPCNGLLRGRCSPSPDSOCFTWELLAMATAE 237  
 Db 176 VLNAYYKQPOEDHALFLFRRAQEVVPGGMVLTILGRSEDRASTRECLLIQOLLAMALNQ 235  
 QY 238 IYSOGCLIDEDKLDTPNIPSYASLEEVKDIYERDGSFTIDHIEGDL-----D 285  
 Db 236 MNSGEIEEKDKDKNTIYOTPTSPEVAEILIEKGSFLIDHASEIYSSCTKDGSGG 295  
 QY 286 SVEMQENDKWKVGEKFTKYVRAATEPTISNQGPELMDKLYDKFTHIVVSOLFAPKPTT 345  
 Db 296 SYEEE-----GYVVARCMRAVAEPILLDHFGAALIEDVFRHYKLLIEMSKETKFI 348  
 QY 346 STILVL 351  
 Db 348 NVIVSL 354

RESULT 4  
 Q9FJZ6 PRELIMINARY: PRT: 354 AA.  
 ID Q9FJZ6  
 AC Q9FJZ6  
 DT 01-MAR-2001 (TREMBlrel. 16, Created)  
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
 DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)  
 DE S-ADENOSYL-L-METHIONININ:SALICYLIC ACID CARBOXYL METHYLTRANSFERASE-LIKE  
 PROTEIN.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;  
 OC Brassicales; Brassicaceae; Arabidopsis.  
 OX NCBI\_TaxID=3702;  
 RN 11  
 RP SEQUENCE FROM N.A.  
 RA STRAIN-COLUMBIA;  
 MEDLINE=98403884; PubMed=9734815;  
 Krcani H., Nakamura Y., Sato S., Asamizu E., Kaneko T., Miyajima N.,









DB 89 -----UCASPAIMLSTMTIDPG--DEVILNPSVSYFTEFAECKIK- 132

QY 271 DASTTIDHTETEDDASVEMOHNNWNGEKEFTVPAFTPIISNGCFEIMKLYDK- 328

DB 133 -----NIDLEDEFNIDFKVKEST-----TKETKLJ-IFNSP--SNPTG-----KYVKEET 175

QY 429 -----FTHLVSD 336

DB 176 IKGIAEIAEDYNLIVSD 193

RESULT 5

FIELD SVDDU STANDARD: PRT: 2185 AA.

AC P13900: Q84794: Q84795: Q84796: Q84797: Q84798: Q84799: Q84800:

AN Q84801: Q84802: Q84803: Q84804:

DB 01-JAN-1990 (Rel. 13, Created)

DB 01-JAN-1990 (Rel. 13, Last sequence update)

DB 30-MAY-2000 (Rel. 39, Last annotation update)

DB GENOME POLYPROTEIN [CONTAINS: COAT PROTEIN VP1 TO VP4; CORE PROTEINS P2A TO P2C; P3A; GENOME-LINKED PROTEIN VP3; PICORNAIN 3C (EC 3.4.22.28) (PROTEASE 3C) (P3C); RNA-DIRECTED RNA POLYMERASE P3D (EC 2.7.7.48)].

DB Swine vesicular disease virus (strain UKC/27/72).

DB Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae; Enterovirus.

DB NCBI\_TaxID:12077;

DB [11]

DB SEQUENCE FROM N.A.

DB MEDLINE-90364770; PubMed-2168111;

DB Szechenyi P., Knowles N.J., McCauley J.W.;

DB "The complete nucleotide sequence of a pathogenic swine vesicular disease virus."

DB Virus Res. 16:255-274(1990).

DB -1- FUNCTION: IT IS THOUGHT THAT THE P2C PROTEIN ATTACHES TO VESICULAR MEMBRANES AND IS ASSOCIATED WITH VIRAL RNA SYNTHESIS.

DB -1- FUNCTION: P3C POLYPEPTIDE IS A PROTEASE THAT CLEAVES AT CERTAIN Q/G SITES IN THE POLYPROTEIN. IT MAY BE A CYSTEINE PROTEASE.

DB -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAHEDRAL UNITS.

DB EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2, VP3, AND VP4.

DB -1- PTM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.

DB -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3

DB This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL collaboration the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DB EMBL: X54521; CA38377.1; -

DB PIR: A30061; GNNYSY.

DB PIR: S11670; S11670.

DB HSSP: P03314; ICQV.

DB MEKOPS: C03.001; -

DB MEKOPS: C03.020; -

DB InterPro: IPR000081; -

DB InterPro: IPR000199; -

DB InterPro: IPR006055; -

DB InterPro: IPR001295; -

DB InterPro: IPR001676; -

DB InterPro: IPR002527; -

DB Pfam: PF00548; Cys-protease-3C; 1.

DB Pfam: PF00947; Pico\_P2A; 1.

DB Pfam: PF01552; Pico\_P2B; 1.

DB Pfam: PF00680; RNA\_dep\_RNA\_pol; 1.

DB Pfam: PF00310; RNA\_helicase; 1.

DB Pfam: PF00073; rhv; 3.

DB Polyprotein; Coat protein; Core protein; Transierase;

DB RNA-directed RNA polymerase; Hydrolase; Thiol protease; Myristate;

DB CHAIN 2 69 COAT PROTEIN VP4 (PIA)

FT CHAIN 70 320

FT CHAIN 331 568

FT CHAIN 569 851

FT CHAIN 852 1001

FT CHAIN 1002 1109

FT CHAIN 1101 1429

FT CHAIN 1430 1518

FT CHAIN 1519 1540

FT CHAIN 1541 1723

FT CHAIN 1724 2185

FT LIPID 2

FT ACT SITE 1687 1687

FT ACT SITE 1701 1701

SO SEQUENCE 2185 AA. 24363 MW. C9H163O322334E1DE CRO64;

Query Match 5.1%; Score 94.5; PB 1; Length 2185;

Best local similarity 19.5%; Fred. No. 20;

Matches 78; Conservative 54; Mismatches 136; Indels 131; Gaps 17;

QY 29 TENAVETLPSRPHLQALNADCGCAACRTTAV-----ISTKKMKK- 72

DB 1420 IEDAV-----VETPGLFAPLDTTSAGYPPYVALGIRKPHLISKRTIDILKLPMDKYG 1872

QY 73 -----KREINCOLDELQVYLMDLFGNDPRLFKG---LSSEV 107

DB 1873 INI PMVTYVNLPLSARKVAKSKSLIFASSI NISAM--GTGDN YKTHPLNPGIVTSSA 1942

QY 108 IGNKEEVPYVMGVESGSHGLFPPNSLHLYNSYVSHMLTQAKGLTSGECLA 162

DB 1933 VG--CD-----PDVEFSKIPVMDGHLIAFDYSGYIASISPMWFTCLKLLKELY 1981

QY 163 LNKGVISKISIPVVEAVLISQHEDF-----IMLANSQVAVNG 205

DB 1982 TNFTNTIF-----YICNSHH YRPYHVFVQZPSCSCSIFNSKTNMILFT 2031

QY 206 GMLVILDG-----KCGSPGSMGSGPVELLAMAIAELVSGGII--DEKILDTN--- 253

DB 2092 LMLVYVGTINLQPMIAVGVNIVASVPMFLAKELALACGYVGLIMTACGCTTNTVT 2091

QY 254 IFSYFASLSE-----VKDLIV-----KCGSTIDILIGTITLSSVPMQDN 292

DB 2092 WYNTVILKPFPEDEQVTVLHFMVMPMPCVHESLEWTFPRKTCGVHSTCLAMWIRGH 2151

QY 293 --EKWVGEKFTKVPATFTITSNCFGEIMKLYDK 329

DB 2152 EYEFKIKIPSVVGPETSLPAPST-----LPPWMLDSF 2185

RESULT 6

ID C81E\_GLYEC STANDARD: PRT: 499 AA.

AC P93147;

DB 15-DEC-1998 (Rel. 37, Created)

DB 15-DEC-1998 (Rel. 37, Last sequence update)

DB 15-DEC-1998 (Rel. 37, Last annotation update)

DB DE CYTOCHROME P450 81E1 (EC 1.14.-.-) (ISOPHAVONE 2'-HYDROXYLASE) (P450 91A4) (CYP GE-3).

DB DE 91A4 (CYP GE-3).

DB CN CYP81E1 OR CYP91A4.

DB OS Glycyrrhiza echinata (Licorice).

DB Eukaryota; Viridiplantae; Umbiferales; Tracheophyta; Spermatophyta;

DB Malvaceae; Glycyrrhizaceae; Glycyrrhizaceae; Glycyrrhizaceae; Glycyrrhizaceae;

DB Fabales; Fabaceae; Papilionoideae; Glycyrrhiza.

DB NCBI\_TaxID:46348;

DB [11]

DB SEQUENCE FROM N.A.

DB Akashi T., Aoki T., Kameya N., Nakamura I., Ayabe S.-I.;

DB "Two new cytochrome P450 cDNAs from alfalfa-induced liver of Glycyrrhiza echinata L." cells."

DB (in) Plant Gene Register PGR97-167.

DB [2]

DB CHARACTERIZATION:

DB MEDLINE-96099233; PubMed-9709098;

KA Akashi T., Aoki T., Ayabe S.-I.:  
 "CyPR181, a cytochrome P450 cDNA of *Ipomoea* (*Cyclopenthera* *echinata*  
 L.), encodes isoflavone 2'-hydroxylase."  
 PL Biochem Biophys Res Commun 261:67-73(1999)  
 CC -1- FUNCTION: CATALYZES THE HYDROXYLATION OF ISOFLAVONES, DAIVERTIN AND  
 FORMONONETIN TO YIELD 2'-HYDROXYISOFLAVONES, 2'-HYDROXYDAIVERTIN,  
 AND 2'-HYDROXYFORMONONETIN. RESPECTIVELY  
 CC -1- PATHWAY: INVOLVED IN THE BIOSYNTHESIS OF ISOFLAVONE-GLYCOSIDES  
 CC -1- ANTIMICROBIAL COMPOUNDS OF LEGUMES.  
 CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC  
 DE EMBL: AB001379; BAA22422.1;  
 DE InterPro: IPR001128;  
 DE Pfam: PF00067; P450; 1  
 DE PRINIS: PR00185; P450;  
 DE PROSITE: PS00086; CYTOCHROME\_P450; 1.  
 KW Oxidoreductase; Monooxygenase; Membrane; Heme.  
 FT BINDING: 436 436 HEME (BY SIMILARITY)  
 FT SEQUENCE: 499 AA: 57191 MW: 47034.6 KDa; 10406 KDa; 10664;  
 SO  
 Query Match 5.1%; Score 93.5; DB 1; Length 499;  
 Best Local Similarity 20.2%; Pred No 3.4;  
 Matches 70; Conservative 56; Mismatches 140; Indels 71; Gaps 17;  
 QY 56 GPNFAVIST---IKRMKRRPFLNGLTFLGVY NLPFNPNTIFKGLSSVIVKNC 112  
 DB 35 GPSLPTITNHLKPLPLAHTFK-----LSKRGHVSLSMPKIVAVSSAS 93  
 QY 113 EEEVY ---VWGVSGHSLKLEFNKSLVLSVSVMLQVAKRGISERIALNK 165  
 DB 84 FFGGFTFKINIVIANRPHLSKYLIVNYITLSTSYDHW--KLPILLIVISNHL 141  
 QY 166 ---GRVYISKTPVDEA---YLSPHEPTMLNPSGFVTVNGVLT---LPFGQ 216  
 DB 142 NFSGIFRFFLQPLTILATSSNFAE---MELSSRIYIMINIMPSIPVYGEQ 198  
 QY 217 SDPDMOSCTWELAAIAIYSGUIDEDKIDFNISY--FASLEE-VMDIVKROS 273  
 DB 199 -VTSIDGASQFDMVSEIIQLISA---NNKTFMPLFLFENLKKPKIJSKITA 253  
 QY 274 F-----TTGIEFTLSEVENGEENFWVEDEKFKVAVETELISNQ 316  
 DB 254 FLFGITFEHPTFRFANTMIHILNQ-ESQPYVITQILKALAMLALHSSAVILF 312  
 QY 317 FG-----PEIMDKLYDKF-THIVVST--EKKLKRTSLIIVISK 353  
 DB 313 WSMNLIHPEVLKRYVDELDTGVGDRIVDESILPKIYIKVNIW 359  
 RESULT 7  
 POLG\_CX3W STANDARD; PRT; 2185 AA.  
 AC Q66282;  
 DT 15-JUL-1999 (Prt. 38, Created)  
 DT 15-JUL-1999 (Rel. 38, last sequence update)  
 DE GENOME POLYPROTEIN [CONTAINS: COAT PROTEIN VP4 (PIB); COAT PROTEIN VP2  
 DE (PIB); COAT PROTEIN VP3 (PIV); COAT PROTEIN VP1 (PII); PROTEIN 2A  
 DE (EC 3.4.22.29) (P2A); COPE PROTEIN P2R; COPE PROTEIN P2C; COPE PROTEIN  
 DE P3A; GENOME-LINKED PROTEIN VP6 (P3B); PICOPAIN 3C (EC 3.4.22.28)  
 DE (PROTASE 3C) (P3C); PNA-DIRECTED PNA POLYMERASE (EC 2.7.7.48) (P3D);  
 DE Coxsackievirus B3 (strain Woodruff).  
 OS Viruses; ssRNA positive-strand viruses, no RNA stage; Picornaviridae;  
 OC Enterovirus.

OX NCBI\_TaxID=103904;  
 RN  
 (1)  
 RP SEQUENCE FROM N.A.  
 PA Knowledge from Team F.S. Infecting F.W. Wessely F. Huber S.G.  
 PT "A mutation in the puff region of VP2 attenuates the myocarditic  
 PT phenotype of an infectious cDNA of the Woodruff virus."  
 PL Submitted (Apr-1999) to the EMBL/GenBank/Trna databases.  
 CC -1- PUFF-108-12A AND THE P2A POLYPEPTIDES ARE PROTEASES THAT CLEAVE  
 CC AT CERTAIN Q/G SITES IN THE POLYPROTEIN. THEY ARE CYSTEINE  
 CC PROTEASES.  
 CC -1- FUNCTION: IT IS THOUGHT THAT THE P2C PROTEIN ATTACHES TO VESICULAR  
 CC MEMBRANES AND IS ASSOCIATED WITH VIRAL RNA SYNTHESIS.  
 CC -1- STRAINT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAHEDRAL UNITS,  
 CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,  
 CC VP3, AND VP4.  
 CC -1- THE SPECIFIC ENZYMATIC ACTIVITIES IN VIVO YIELD MATUR PROTEINS.  
 CC CLEAVAGE BETWEEN VP4 AND VP2 IS ANTICATALYZED; VP1/P2A IS  
 CC CATALYZED BY P2A; ALL OTHER CLEAVAGES ARE CATALYZED BY P2C.  
 CC -1- SIMILARITY: P2A PROTEASE BELONGS TO PEPTIDASE FAMILY C3.  
 CC -1- SIMILARITY: P2C PROTEASE BELONGS TO PEPTIDASE FAMILY C3.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC  
 DE EMBL: U52956; AAP02228.1;  
 DE MEGPUS: G03.011;  
 DE MEGPUS: G03.022;  
 DE InterPro: IPR000081;  
 DE InterPro: IPR000199;  
 DE InterPro: IPR006605;  
 DE InterPro: IPR01205;  
 DE InterPro: IPR001643;  
 DE InterPro: IPR001676;  
 DE InterPro: IPR002527;  
 DE Pfam: PF00548; Cys-Protease-3C; 1.  
 DE Pfam: PF00447; Pico\_P2A; 1.  
 DE Pfam: PF00472; Pico\_P2B; 1.  
 DE Pfam: PF00680; PNA-DIR\_PNA\_P1; 1  
 DE Pfam: PF00640; PNA-Helase; 1.  
 DE Pfam: PF00073; Hiv\_3  
 DE PRINIS: PR00918; CALICIVIRUS.  
 KW Polypeptide; Coat Protein; Hydrolyase; Thiol Protease;  
 KW PNA-directed PNA polymerase; Hydrolyase; Thiol Protease;  
 FT CHAIN 2 69  
 FT CHAIN 70 332  
 FT CHAIN 333 579  
 FT CHAIN 571 851  
 FT CHAIN 852 1001  
 FT CHAIN 1002 1100  
 FT CHAIN 1101 1429  
 FT CHAIN 1430 1519  
 FT CHAIN 1519 1540  
 FT CHAIN 1541 1723  
 FT CHAIN 1724 2185  
 FT LIPID 2  
 FT ACT\_SITE 1687 1687  
 FT ACT\_SITE 1701 1701  
 FT ACT\_SITE 1701 1701  
 FT SEQUENCE 2185 AA: 24369 MW: 11934.67 KDa; 25294 KDa; 10664;  
 SO  
 Query Match 5.1%; Score 93.5; DB 1; Length 2185;  
 Best Local Similarity 20.2%; Pred No 2.4;  
 Matches 69; Conservative 134; Mismatches 134; Indels 134; Gaps 18;  
 QY 29 LKNAVTFSTHEQAINNAAGGCAAGNTFAV-----ISTIKRMEEV-- 72  
 DB 1920 LKNAV-----VETPKLPAIDLTGNGPYVALGIKKKPIHLSKTPRLTKIKRMKNGV 1872

```

67 73 -----KGRFLNCGTLEIQQVYINLFGNDNTLFGK---LSSEV 107
68 1873 LNLPMVTVVKRPFSAEKVAKGKSPILKASLNTSVAMPQTFNKKTFPHNPPVAVTQSA 1932
69 108 IGNKCEVPCVYMGVPGSFHGRLEPRNSLHLVHSSVSMITQAP-----KGLTSRKL 161
70 1943 VG--CD-----PDLFMSKIPVMDLGHLLADYSGVDASLSPWFACLLKLEKLY 1981
71 162 ALNFKGVYISKTSFVVFVAVLSQFEDP-----TMLNASSQVVRN 204
72 1982 S-HKEINVID-----YLNSSHLYPKKHVFPVQMPSCSCSTISFNSMNNILK 2330
73 205 GGMVYILP-----PGSDPDMQSGFTWELAMALVEL-VSQGLT--DECDLTFPN 253
74 2041 TMLKVKYKIDLDQFPMIAYRNVVASYPPMPTASLLAEAGKRYGLIMTPAKGEPFNEV 2090
75 254 -----IPSYFASLE-----VKDIVE-----KQASFTLDHESFELDSVEMQ 291
76 2094 TWIVNTPFKRPPVAFPPFVHFVPMKCHESIPWFKIKRNGQHVSLCLLAWHND 2150
77 292 N--LRKWPCEKFTKVVAFPPFISNPGFELMKKLYKK 329
78 2151 HEYEFIFKIRSVVPCGLTLPAPST-----IPRKMILSF 2185

```

## RESULT 8

P014\_CXA9

P014\_CXA9 STANDARD: PRT: 2201 AA.

```

67 01-MAY-1991 (Ref. 18, Created)
68 01-AUG-1991 (Ref. 19, Last sequence update)
69 30-MAY-2000 (Ref. 19, Last annotation update)
70 GENOME POLYPEPTIDE [CONTAINS: COAT PROTEIN VP4 (P1A); COAT PROTEIN VP2
71 (P1B); COAT PROTEIN VP3 (P1C); COAT PROTEIN VP1 (P1D); CORE PROTEIN
72 P2A; CORE PROTEIN P2B; CORE PROTEIN P2C; CORE PROTEIN P3A; GENOME-
73 LINKED PROTEIN VP6 (P3B); PICORNAIN 3C (EC 3.4.22.28) (PROTEASE 3C)
74 (P3C); RNA-DIRECTED RNA POLYMERASE (EC 2.7.7.48) (P3D)].
75 COXSACKIEVIRUS A9 (strain Grijps).
76 VIRUSES; ASPMA positive-strand viruses, no RNA stage; Picornaviridae;
77 Enterovirus.
78 NCBI_TaxID=12068;
79 [1]
80 SEQUENCE FROM N.A.
81 MEDLINE=90111704; PubMed=2556158;
82 Chang K.H., Avnion P., Hyypia T., Stanway G.,
83 "The nucleotide sequence of coxsackievirus A9: implications for
84 receptor binding and enterovirus classification."
85 J. Gen. Virol. 70:3269-3280(1989).
86 [2]
87 X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS) OF 1-870.
88 MEDLINE=20114480; PubMed=10647183;
89 Hendry E., Hatanaka H., Fry R., Smyth M., Tate J., Stanway G.,
90 Sautli J., Maroney M., Hyypia T., Stuart D.:
91 "The crystal structure of coxsackievirus A9: new insights into the
92 uncoupling mechanisms of enteroviruses."
93 Structure 7:1527-1538(1999).
94 -1- FUNCTION: IT IS THOUGHT THAT THE P2C PROTEIN ATTACHES TO VESICULAR
95 MEMBRANES AND IS ASSOCIATED WITH VIRAL RNA SYNTHESIS.
96 -1- FUNCTION: P3C POLYPEPTIDE IS A PROTEASE THAT CLEAVES AT CERTAIN
97 Q/G SITES IN THE POLYPEPTIDE. IT MAY BE A CYSTEINE PROTEASE.
98 -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF NO. 10, SAHABAT UNITS,
99 EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
100 VP3, AND VP4.
101 -1- PPM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MAJORE PROTEINS.
102 -1- SIMILARITY: THE PROTEASE REFONES TO PROTEASE FAMILY C3.

```

```

67 This SWISS-PROT entry is copyright. It is provided through a collaboration
68 between the Swiss Institute of Bioinformatics and the EMBL outstation
69 at the European Bioinformatics Institute. There are no restrictions on its
70 use by non-profit institutions as long as its content is in no way
71 modified and this statement is not removed. Usage by and for commercial
72 entities requires a license agreement (See: http://www.isb-sib.ch/announcement/
73 or send an email to license@isb-sib.ch).

```

```

67 CC EMBL: D00627; BA00518.1;
68 PIP: Q0553; GNNVAA9.
69 DB: 1D4M; 23-DEC-99.
70 DR MEROPS: C03.011;
71 DR MEROPS: C03.022;
72 DR InterPro: IPR000081;
73 DR InterPro: IPR000199;
74 DR InterPro: IPR000625;
75 DR InterPro: IPR001205;
76 DR InterPro: IPR01676;
77 DR InterPro: IPR02527;
78 DR Ffam: PF00548; Cys-Protease-3C; 1.
79 DR Ffam: PF00947; Pico_P2A; 1.
80 DR Ffam: PF01552; Pico_P2B; 1.
81 DR Ffam: PF00680; RNA_dep_RNA_pol; 1.
82 DR Ffam: PF00910; RNA_helicase; 1.
83 DR Ffam: PF00773; Thy; 3.
84 DR Polyprotein; Coat protein; Core protein; Hydrolyase; Thiol protease; Myristate.
85 KW RNA-directed RNA polymerase; Hydrolyase; Thiol protease; Myristate.
86 FT CHAIN 2 69
87 FT CHAIN 70 330
88 FT CHAIN 331 568
89 FT CHAIN 569 870
90 FT CHAIN 871 1017
91 FT CHAIN 1018 1116
92 FT CHAIN 1117 1445
93 FT CHAIN 1446 1534
94 FT CHAIN 1535 1556
95 FT CHAIN 1557 1739
96 FT CHAIN 1740 2201
97 FT LIPID 2
98 FT ACT_SITE 1703 1703
99 FT ACT_SITE 1717 1717
00 FT SITE 858 860
01 SEQUENCE 2201 AA; 24633 MW; 22EAGF3EBE336CF GECK4;

```

Query Match 5.08; Score 92.5; Pos 1; Length 2201;

Host Local Similarity 19.88; Prod No. 29; Matches 79; Conservative 55; Mismatches 133; Indels 133; Gaps 18;

```

67 29 LENVETLFSPDEPHLOALNADLQAGCPTFAV-----ISTIKRMK-- 72
68 1836 LEDAV-----YCFLEALDILISAGDYVALGKIKKILSKILDLIKLBMQKYG 1888
69 73 -----KGRFLNCGTLEIQQVYINLFGNDNTLFGK---LSSEV 107
70 1889 LNLPMVTVVKRPFSAEKVAKGKSPILKASLNTSVAMPQTFNKKTFPHNPPVAVTQSA 1948
71 108 IGNKCEVPCVYMGVPGSFHGRLEPRNSLHLVHSSVSMITQAP-----KGLTSRKL 161
72 1949 VG--CD-----PDLFMSKIPVMDLGHLLADYSGVDASLSPWFACLLKLEKLY 1997
73 162 ALNFKGVYISKTSFVVFVAVLSQFEDP-----TMLNASSQVVRN 204
74 1998 S-HKEINVID-----YLNSSHLYPKKHVFPVQMPSCSCSTISFNSMNNILK 2046
75 205 GGMVYILP-----PGSDPDMQSGFTWELAMALVEL-VSQGLT--DECDLTFPN 253
76 2041 TMLKVKYKIDLDQFPMIAYRNVVASYPPMPTASLLAEAGKRYGLIMTPAKGEPFNEV 2106
77 254 -----IPSYFASLE-----VKDIVE-----KQASFTLDHESFELDSVEMQ 291
78 2107 TWIVNTPFKRPPVAFPPFVHFVPMKCHESIPWFKIKRNGQHVSLCLLAWHND 2166
79 292 N--LRKWPCEKFTKVVAFPPFISNPGFELMKKLYKK 329
80 2167 HEYEFIFKIRSVVPCGLTLPAPST-----IPRKMILSF 2201

```

RESULT 9  
EF5\_MOUSE

ID	EFZF5_MOUSE	STANDARD	ENT	335 AA
AC	061502;			
DT	01-NOV-1997 (Ref. 35, Created)			
DT	01-NOV-1997 (Ref. 35, last sequence update)			
DT	30-MAY-2006 (Ref. 39, last annotation update)			
DE	TRANSCRIPTION FACTOR EFZF5 (EFZF-5).			
GN	EFZF5.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Cladibia; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
PN	NCBI_TaxID=10090;			
FX	[1]_TaxID=10090;			
PP	SEQUENCE FROM N A			
PX	MEDLINE=9349934; PubMed=7542769;			
RA	Buck V., Allen K.B., Soerensen T., Ryboe A., Hifmans E.M.,			
RA	Wolthuis-van't Hof M., Reinders F., Ja Hengue N.R.;			
RT	"Molecular and functional characterisation of EFZF5, a new member of			
RT	the EFZF family.";			
RL	Oncogene 11:31-38(1995).			
RN	[2]			
RP	DEVELOPMENTAL EXPRESSION.			
RX	MEDLINE=948025478; PubMed=9476416;			
RA	Dagnino L., Fry C.J., Bartley S.M., Farnham P., Gallie B.L.,			
RA	Phillips F.A.;			
RT	"Expression patterns of the EFZF family of transcription factors during			
RT	mouse development system development.";			
RL	Mech. Dev. 66:13-25(1997).			
RN	[3]			
RP	DEVELOPMENTAL EXPRESSION.			
PX	MEDLINE=97234322; PubMed=9149906;			
RA	Dagnino L., Fry C.J., Bartley S.M., Farnham P., Gallie B.L.,			
RA	Phillips F.A.;			
RT	"Expression patterns of the EFZF family of transcription factors			
RT	during murine epithelial development.";			
RL	Cell Growth Differ. 8:553-563(1997).			
CC	-1- FUNCTION: TRANSCRIPTIONAL ACTIVATOR THAT BINDS TO EFZF SITES. THESE			
CC	SITES ARE PRESENT IN THE PROMOTER OF MANY GENES WHOSE PROMOTERS ARE			
CC	INVOLVED IN CELL PROLIFERATION. MAY MEDIATE GROWTH FACTOR-			
CC	-1- INITIATED SIGNAL TRANSDUCTION.			
CC	-1- SUPRINITY: COMPONENT OF THE DRP1/EFZF TRANSCRIPTION FACTOR COMPLEX.			
CC	BINDS COOPERATIVELY WITH DRP-1 TO EFZF SITES. INTERACTS			
CC	PREFERENTIALLY WITH RETINOLACTAM-RELATED PROTEIN PROTEINS P130			
CC	AND TO A LESSER EXTENT WITH P107, THAT INHIBIT THE EFZF			
CC	TRANSACTIVATION DOMAIN.			
CC	-1- SUBCELLULAR LOCATION: NUCLEAR.			
CC	-1- DEVELOPMENTAL STAGE: IN THE DEVELOPING EPIDERMIS, FIRST DETECTED			
CC	IN 13.5-14.5 DPC EMBRYOS. WITH THE APPEARANCE OF STRATIFIED			
CC	EPITHELIUM, LEVELS OF EFZF-5 EXPRESSION INCREASE AND BY 16.5 DPC,			
CC	HIGH EXPRESSION POINT IN THE SUPRAVASCULAR CELL LAYERS. HIGH			
CC	EXPRESSION ALSO FOUND IN OTHER REGIONS WITH STRATIFIED SQUAMOUS			
CC	EPITHELIA INCLUDING THE DEVELOPING PALATE, LIP AND TONGUE IN THE			
CC	DEVELOPING NERVOUS SYSTEM. FIRST DETECTED IN THE FOREBRAIN AT 9.5			
CC	DPC. AT DAY 10.5, STRONGLY EXPRESSED IN THE POSTAL REGION OF THE			
CC	SPINAL CORD. BY DAY 11.5, EFZF-5 IS EXPRESSED THROUGHOUT THE			
CC	DEVELOPING CENTRAL NERVOUS SYSTEM. IN 12.5-15.5 DPC EMBRYOS,			
CC	EXPRESSION FOUND IN THE UNDIFFERENTIATED VENTRICULAR REGIONS OF			
CC	THE BRAIN. IN THE RETINA, EXPRESSED IN 14.5-18.5 DPC EMBRYOS. IN			
CC	THE RETINOLACTIC CELL LAYER IN OTHER DEVELOPING TISSUES, HIGHLY			
CC	EXPRESSED IN THE CHOROID PLEXUS. ALSO FOUND IN THE KIDNEY, LIVER,			
CC	LUNG, HEART AND WEAKLY IN DEVELOPING SKELETAL MUSCLE AND			
CC	CHONDROCYTES.			
CC	-1- SIMILARITY: BELONGS TO THE EFZF/DP FAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed, changed by and for commercial			
CC	entities requires a license agreement (See <a href="http://www.isb.slb.ch/submitting/">http://www.isb.slb.ch/submitting/</a>			
CC	or send an email to: <a href="mailto:license@isb.slb.ch">license@isb.slb.ch</a> )			
CC	-----			
OR	EMBL; X86925, CAA65608.1,			
OR	MGD; MG1:105091; E2F5.			

Query Match: 5.0%; Score 92; dp: 1; Length 335;  
Best Local Similarity: 23.5%; Pred. No. 2.6; Mismatches 135; Indels 86; Gaps 15;  
Matches 79; Conservative 36;

FW	Transcription regulation; Activator, DNA-binding, Nuclear protein.
FT	DNA_BIND 37 108 POTENTIAL.
FT	DOMAIN 66 88 LEUCINE ZIPPER.
FT	DOMAIN 71 108 DEF BOX.
FT	DOMAIN 109 205 DIMERIZATION (POTENTIAL).
FT	DOMAIN 277 335 TRANSCAIVATION (POTENTIAL).
FT	DOMAIN 312 329 P130 PROTEIN ASSOCIATION (POTENTIAL).
FT	DOMAIN 223 226 POLY-SER.
SW	SEQUENCE 445 AA; 46784 MW; EAPVVD5HAFQIPE2 (P07544);

Query Match: 5.0%; Score 92; dp: 1; Length 335;  
Best Local Similarity: 23.5%; Pred. No. 2.6; Mismatches 135; Indels 86; Gaps 15;  
Matches 79; Conservative 36;

FW	Transcription regulation; Activator, DNA-binding, Nuclear protein.
FT	DNA_BIND 37 108 POTENTIAL.
FT	DOMAIN 66 88 LEUCINE ZIPPER.
FT	DOMAIN 71 108 DEF BOX.
FT	DOMAIN 109 205 DIMERIZATION (POTENTIAL).
FT	DOMAIN 277 335 TRANSCAIVATION (POTENTIAL).
FT	DOMAIN 312 329 P130 PROTEIN ASSOCIATION (POTENTIAL).
FT	DOMAIN 223 226 POLY-SER.
SW	SEQUENCE 445 AA; 46784 MW; EAPVVD5HAFQIPE2 (P07544);







```

FT CHAIN 2 69 COAT PROTEIN VP4 (PIA).
FT CHAIN 70 330 COAT PROTEIN VP2 (PIB).
FT CHAIN 341 568 COAT PROTEIN VP3 (PIC).
FT CHAIN 569 851 COAT PROTEIN VP1 (PID).
FT CHAIN 852 1001 GAGE PROTEIN P2A (P2.1B).
FT CHAIN 1002 1100 GAGE PROTEIN P2B (P2.2B).
FT CHAIN 1101 1429 GAGE PROTEIN P2C (P2.3B).
FT CHAIN 1430 1518 GAGE PROTEIN P2A (P2.1B).
FT CHAIN 1519 1540 GAGE PROTEIN P2B (P2.2B).
FT CHAIN 1541 1723 GAGE PROTEIN P2C (P2.3B).
FT CHAIN 1724 2183 RNA-DEPENDENT RNA POLYMERASE P40.
FT LIPID 2 MYRISTATE (BY SIMILARITY).
FT ACT_SITE 1687 1687 PROTEASE (POTENTIAL).
FT ACT_SITE 1701 1701 PROTEASE (POTENTIAL).
SQ SEQUENCE 2185 AA; 241164 MW; 5P2295C094C0B3B6 C0C04.

Query Match 5.0%; Score 91.5; DB 1; Length 2185.
Best Local Similarity 19.5%; Pred. No. 35;
Matches 78; Conservative 53; Mismatches 137; Indels 141; Gaps 17.

UY 29 LENVETLEPSKSHLQALNADGZAGNTFAV ..... ISTIKRMER.. 72
UY 1820 LERAV .....
UY 73 ..... KRELNCQLELYNDLPGDNFLFG---LSSEV 107
UY 1873 LNPMTVYVKTELRSADKVKAKKSPLEASSINDSAMPTFCNLKFRPH NPGVITCSA 1922
UY 108 LKCKCEVPYVWGVCSGSHGLEPPNSLHIVSSVYVMTQAKRLTREGCA----- 162
UY 1933 VG---CD-----PDVWWSKLPVMDLHJAFDSYDASLSVWFLTKLILEKGY 1981
UY 163 LNKRIYISKTSPPVREAVLSQHEDE-----TWLNARSQEVVNG 205
UY 1982 TNEFTNID-----YLNSHILYEDKRYVWGMSSGSSNIFNSHINILLFT 2033
UY 206 GCVILLIG-----KCSNPDSMGSCTWELLAMATFL-VSQCIL--DEKIDTFN--- 253
UY 2032 LMKVVKGIDCLQFMFMAVGDDVLAIVPILASLAEAGSGYGLIMTACKGCEVNEVT 2091
UY 254 -----IPSYFASIFP-----VRLIVE-----PMSFPIHIEPDSVSENGEN 292
UY 2092 WIVNTLEKRFKRAFDQYPTLVHVPVMKDIHSIKTKPKPTQCHVSLCLAHNGEH 2151
UY 293 --DKWVGEKFTKVRATFEPILISMGPELMUKLYDK 329
UY 2152 EYELFKIKLKVAVVCEGLSIFAST ...LEKWLDSF 2185

RESULT 13
P0R2_CHITE STANDARD; PRT; 1371 AA.
AC 026255;
DE 01-NOV-1997 (Ref. 35, created)
DE 01-NOV-1997 (Ref. 35, last sequence update)
DE 30-MAY-2000 (Ref. 39, last annotation update)
DE TRIFUNCTIONAL PURINE BIOSYNTHETIC PROTEIN ADENOSINE-3 [INCLUDES:
DE PHOSPHORIBOSYLAMINE--GLYCINE LIGASE (EC 6.3.4.13) (GARS) (GLYCINAMIDE
DE RIBONUCLEOTIDE SYNTHETASE) (PHOSPHORIBOSYLGLYCINAMIDE SYNTHETASE);
DE PHOSPHORIBOSYLGLYCINAMIDE CYTIDYL-TLASE (EC 6.3.3.1) (AIRS)
DE (PHOSPHORIBOSYL-AMINOIMIDAZOLE SYNTHETASE) (AIR SYNTHASE);
DE PHOSPHORIBOSYLGLYCINAMIDE FORMYLTANSFERASE (EC 2.1.2.2) (GART) (GAR
DE TRANSFORMYLASE) (5'-PHOSPHORIBOSYLGLYCINAMIDE TRANSFORMYLASE)]
GN GART.
OS Chironomus tentans (Midge).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Plecoptera; Neoptera; Emdopterygota; Diptera; Nemertocera;
OC Chironomidae; Chironomidae; Chironomidae; Chironomus.
OX NBI_TaxID=7153;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92489368; PubMed=1518084;

```

```

RA Clark D.V., Benikoff S.;
RT "Unusual organizational features of the Drosophila Gart locus are not
RT conserved within Diptera."
PI J. Mol. Evol. 35:51-59(1992).
CC 1 CATALYTIC ACTIVITY: ATP + 5'-PHOSPHORIBOSYLAMINE + GLYCINE -> ADP +
CC ORTHOPHOSPHATE + 5'-PHOSPHORIBOSYLGLYCINAMIDE.
CC 1 CATALYTIC ACTIVITY: 10-FORMYLETANUORICATE + 5'-PHOSPHO-
CC RIBOSYLGLYCINAMIDE -> TETRAHYDROFOLATE + 5'-PHOSPHORIBOSYL-
CC N-FORMYLTGLYCINAMIDE.
CC 1 CATALYTIC ACTIVITY: ATP + 5'-PHOSPHORIBOSYL FORMYLTGLYCINAMIDE
CC ADP + ORTHOPHOSPHATE + 5'-PHOSPHORIBOSYL 5' AMINOTRIMETAZOLE.
CC 1 PATHWAY: SECOND, THIRD AND FIFTH STEPS IN DE NOVO PURINE
CC BIOSYNTHESIS.
CC 1 SIMILARITY: IN THE N-TERMINAL SECTION, BELONGS TO THE GARS FAMILY.
CC 1 SIMILARITY: TO OTHER AIRS AND GART FROM BAUTERIA AND EUKARYOTES.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use, by non-specific institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb.scrib.ch).
CC or send an email to license@isb.scrib.ch).
CC EMBL: S43653; AAB23115.1;
CC DR HSSP: P08179; IGAR.
CC DR InterPro: IPR000115;
CC DR InterPro: IPR000728;
CC DR InterPro: IPR001555;
CC DR InterPro: IPR002376;
CC DR Pfam: PF00586; AIRS; 2;
CC DR Pfam: PF01071; GARS; 1;
CC DR Pfam: PF00591; formyl_transf; 1;
CC DR PROSITE: PS00184; GARS; 1;
CC DR PROSITE: PS00733; GART; 1;
CC DR Multifunctional: enzyme, Purine biosynthesis, ligase, transferase.
CC FT DOMAIN 1 433 GARS.
CC FT DOMAIN 434 1171 AIRS.
CC FT DOMAIN 1172 1315 GART.
CC FT ACT_SITE 1315 1315 BY SIMILARITY.
CC SQ SEQUENCE 1371 AA; 149103 MW; 7046640d153b0446 C0C04.

Query Match 4.9%; Score 91; DB 1; Length 1371;
Best Local Similarity 18.9%; Pred. No. 21;
Matches 66; Conservative 56; Mismatches 142; Indels 96; Gaps 17.

UY 41 VASMAIPALFNAPVIFSSPHHLALNADGZAGNTFAV.....ISTIKRMER.. 72
UY 529 VAMCVNIVLCNAGAPIGELY.....IACGHLEVPVAVIVAGIDGRKANCA 577
UY 80 ----QTLEQVYLN-----DLFG-----NDFNTIFKLSSEVYGNKCEVPCVMAVGS- 125
UY 578 LGGSTAFAMSMYGRKRYLACYGAGIIFYEELIPALNDVHVG-----VIGGLPSSG 610
UY 126 FH-----GRLEPPNSLHV-----HSSYVAMLTQAPKILTSFENALAKKIYI 170
UY 631 IHSNGFSVINKVIFQOIGFKLIDIAEPSDSIKSYGMEFLTP-----RLVY 675
UY 171 SKISPVVEATLSQHEDETHLNAHSQEVVNGVWLLIPKQCSQSSIMSGCTWEL 210
UY 676 SPTT-PTFNQGVKALAMITGGGLFPTPTILTRHISQCI .....DALTWK 720
UY 231 LAMAVELVSQGLID EDRLTFN.....HSTVASLPEVKS .....YDEKSTPT 725
UY 721 LKPVFSVLAAGHVANANENLPTFNGCGMIIIMPNDIEMPTIDEAVMISITQDRHNGP 780
UY 276 IADLEELCLSSVPMGKLNKKNVCKLKVAVPAVDFITLSNCTHLEKMKK 325
UY 781 QVIVKNP--KEVLEKREVIHMKKGDAELISI-SYDSQVDITAGNELVNI 827

RESULT 14

```



CC VP3, AND VP4.  
 CC -1- PPM: SPECIFIC ENZYMATIC CLEAVAGES, IN VIVO YIELD MATURE PROTEINS.  
 CC CLEAVAGE BETWEEN VP4 AND VP2 IS AUTOCATALYTIC; VP1/PPA IS  
 CC CATALYZED BY P2A. ALL OTHER CLEAVAGES ARE CATALYZED BY P3C  
 CC -1- SIMILARITY: P2A PROTEASE BELONGS TO PEPTIDASE FAMILY 03.  
 CC -1- SIMILARITY: P3C PROTEASE BELONGS TO PEPTIDASE FAMILY 03.  
 CC This SWISS-PROT entry is copyright © it is provided through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation.  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DB EMBL: X80059; CAA56465.1; -  
 DB EMBL: D10582; BAA0149.1; -  
 DB PIR: A36642; GNNVEC.  
 DB MEROPS: C03.001; -  
 DB InterPro: IPR000081; -  
 DB InterPro: IPR000199; -  
 DB InterPro: IPR000605; -  
 DB InterPro: IPR001203; -  
 DB InterPro: IPR001643; -  
 DB InterPro: IPR001676; -  
 DB InterPro: IPR002527; -  
 DB Pfam: PF00548; Cys-protease-3C; 1.  
 DB Pfam: PF00947; Pico\_P2A; 1.  
 DB Pfam: PF01552; Pico\_P2B; 1.  
 DB Pfam: PF00680; RNA\_dep\_RNA\_pol; 1.  
 DB Pfam: PF00910; RNA\_helicase; 1.  
 DB Pfam: PF00073; hmv; 3.  
 DB PRINTS: PRO0918; CALICIVIRUSNS.  
 KW Polypeptide; Coat protein; Core protein; Myristate;  
 KW RNA-directed RNA polymerase; Hydrolase; Thiol protease.  
 FT CHAIN 2 69  
 FT CHAIN 70 331  
 FT CHAIN 332 569  
 FT CHAIN 570 861  
 FT CHAIN 862 1011  
 FT CHAIN 1012 1110  
 FT CHAIN 1111 1438  
 FT CHAIN 1440 1528  
 FT CHAIN 1529 1550  
 FT CHAIN 1551 1733  
 FT CHAIN 1734 2195  
 FT LIPID 2 2  
 FT ACT\_SITE 1697 1697  
 FT ACT\_SITE 1711 1711  
 FT CONFLICT 823 827  
 SQ SEQUENCE 2195 AA; 245405 MW; 10CF5DF288831AF0 CRC64;

Query Match 4.9%; Score 90.5; DR 1; Length 2195;  
 Best Local Similarity 19.8%; Pred. No. 43;  
 Matches 79; Conservative 53; Mismatches 135; Indels 133; Gaps 18;

QY 29 LENAFTLPSRPHQAI NNAIDGCAACNTFAV-----ISTIKPMMEK-- 72  
 DB 1830 LENAFTLPSRPHQAI NNAIDGCAACNTFAV-----ISTIKPMMEK-- 1832  
 QY 73 -----KREINCOULELOVYLNILFGNDNTLEKG---LSSEV 107  
 DB 1883 LNLPMYVYKDELPSALVKAKGKSPILFASSI NISVAMPPTPLNLPJIVTCSA 1942  
 QY 108 LCKKCEVPPYVWGVPQSEFGRLFPKNSILHVSYSVHMLIQAP-----KGLTSREGI 161  
 DB 1943 VG--CG-----PDLFWSKIPLYMLDGHLLAFDYSGLDASLPVWACILKILLEKGY 1991  
 QY 162 ALNKGKLYISKTSPPVVPFAVI SQPHFPP-----TMLNAPSGQEVVFN 204  
 DB 1992 L HKENYID-----YLGNSHHLYPDHNYFFPGGMPGSGYSQISMFNENINNIITR 2040

QY 205 GCVVILIRG-----EGCSPPSGMSQFTWELAMAIAGL-VSQDIL--DEKRLDTEFN-- 253  
 DB 2041 TIMKVKYKQIDLQCFRMLAVGNVIVASYPMDIASIPACTCKQYGLMTPAKKSGTNEV 2100  
 QY 254 - IFSYFALEDE - VKDLYE---EGCSPTDRIHRTPLDSVEMQE 291  
 DB 2101 LKINVLKRYFALEDEYFGLVHIVKQKMLIHESLWILFIRNILEGKVSQTLAMHNGE 2169  
 QY 292 N--LKNVKGKELTKVVPATPELISNQPCELMKCYKPK 329  
 DB 2161 HEVEFFPKTRSPVVGPCI TLPAPST-----LPPWINDSF 2195

Search completed: August 31, 2001, 15:07:34  
 Job time: 372 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Computer Ltd

OM protein - protein search, using sw model

Run on: August 31, 2001, 14:25:47 : Search time 38.18 seconds  
(without alignments)  
710.271 Million cell updates/sec

Title: US-09-577-657A-1  
Perfect score: 1847  
Sequence: 1 FMNNGEESSTVAQNSSTQQ LEAKPKTSTIIIVISKIIS 356

Scoring table:  
BLOSUM62  
Gapop 10 0, Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :  
1: p1f68:\*  
2: p1f2:\*  
3: p1f3:\*  
4: p1f4:\*

Pred No is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	630	34.1	371	2 E85430	hypothetical prote
2	570	30.9	359	2 E84513	hypothetical prote
3	504.5	27.3	363	2 E96703	hypothetical prote
4	499.5	27.0	351	2 A86285	protein P91.6 [im
5	496.5	26.9	348	2 T51783	Atpp-like protein
6	485.5	26.3	348	2 T51781	protein kinase ACP
7	480	26.0	379	2 T51784	Atpp-like protein
8	480	26.0	519	2 T05052	hypothetical prote
9	475.5	25.7	323	2 B86329	hypothetical prote
10	468	25.7	340	2 E84309	hypothetical prote
11	468	25.7	340	2 E84309	hypothetical prote
12	99	5.4	1016	2 T15137	hypothetical prote
13	97.5	5.3	458	2 T10506	hypothetical prote
14	97	5.3	1696	2 T33617	ribulose-bisphosph
15	95.5	5.2	2193	2 E44194	hypothetical prote
16	95.5	5.2	2193	2 E44194	hypothetical prote
17	94.5	5.1	375	2 A64300	aspartate transami
18	94.5	5.1	787	2 H71453	hypothetical prote
19	94.5	5.1	2185	1 GNNVSV	genome polyprotein
20	94	5.1	983	2 G84524	probable disease t
21	93.5	5.1	525	2 T41663	probable transcript
22	93	5.0	719	2 E23467	probable transcript
23	93	5.0	885	2 E47533	glucose 6 phosphat
24	92.5	5.0	2281	1 GNNVAV	glucose 6 phosphat
25	92	5.0	335	2 T43652	E2P-5, mouse
26	92	5.0	342	2 T43652	hypothetical prote
27	91.5	5.0	1308	2 T15280	hypothetical prote
28	91.5	5.0	1820	2 E83207	conserved hypofact
29	91.5	5.0	2185	1 GNNVBI	genome polyprotein

30	91.5	5.0	2185	1 GNNVSH	genome polyprotein
31	91	4.9	550	2 F01770	hypothetical prote
32	91	4.9	2182	1 GNNVBI	genome polyprotein
33	90.5	4.9	1374	1 GNNVFC	genome polyprotein
34	90	4.9	588	2 T27443	hypothetical prote
35	90	4.9	664	2 T29011	hypothetical prote
36	90	4.9	1123	2 A72311	hypothetical prote
37	90	4.9	2231	2 E71870	conserved hypofact
38	90	4.9	2304	2 T07920	hypothetical prote
39	89.5	4.8	568	2 F71614	probable acetyl-Co
40	89.5	4.8	1032	2 S53571	chromatinic RING t
41	89.5	4.8	1192	2 S55100	hypothetical prote
42	89.5	4.8	2193	2 S53919	hypothetical prote
43	89	4.8	562	2 S27800	polypeptide (1A, 1
44	89	4.8	570	2 T32061	elastase precursor
45	89	4.8	676	2 S41022	gut-specific carbo
					hypothetical prote

## ALIGNMENTS

RESULT 1  
E85430  
hypothetical protein A436473 [unc-rted] Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 16-Feb-2001 #sequence-revision 16-Feb-2001 #text-change 16-Feb-2001  
C:Accession: E85430  
R:anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Sp  
Nature 402, 769-777, 1999  
A:Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.  
A:Cross references: DB:NC\_001068, NID:97273995, E:08:AB0013, E:3508:G00140  
A:Accession: E85430  
A:Reference number: A85001, MIMD:20083488  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1371 <STOP>  
A:Cross references: DB:NC\_001068, NID:97273995, E:08:AB0013, E:3508:G00140  
C:Genetics:  
A:Gene: AT9G36470  
A:Map position: 4

Query Match	34.1%	Score 630	DB 2	Length 371
Best Local Similarity	37.3%	Pred No 5.4e+43		
Matches 135	Conservative 84	Mismatches 131	Indels 12	Gaps 5
QY	1	FMNNGEESSTVAQNSSTQQ	MAAPALBNVETTESKAFHLMALNADGAAAPNIF	60
DB	11	YMTGSGKTSVAANSSTQKKASPTAKHITLFLQGLYKEIFPKSLADDSGSPWIL		69
QY	61	AVIST-1KMKKKKPELNGTE-ELQVYLNLFNNENITFKGISSEVIGK		111
DB	76	SLIHLKIVGVNHHRIHIGLEPESFINDPGNDFEIKSLDPIHIEKIKANNCD		129
QY	112	CEVNVVYVWVPSFTELETKNSLRVHSYSVIMLZAEKELTEKLALNKKKIVIS		171
DB	130	CPSPV-FLAAYPSFGRLPEPNTIFHYVASHLSMLSVPAIDEGKSLNKKCVSIC		187
QY	172	KTSIVVFVAVYSGFRTGTMILNASEVNVTRGCVVLTAKNQSLSKSKSPTWILL		231
DB	188	STSESAVSKAVTSQFKHDSIFLPCSTFEMVSAIMVLLTSPFGVTFENSHFWILL		247
QY	232	AMALIVVGGIIDEKRTIFNTFTFASLEFVKIVHPSSTFTTCTCTTCSEVKE		291
DB	248	SRKAAVAVAGTTEFKLISYIMHFAFSAITFESVAKKSGFTPEIPMLFVKKDKNT		307
QY	272	NEAWVGEENTFVAVNLDTLSNDFEIMQLYDKLHIVSOLAKLNLSILLV		351
DB	306	ESGVVAKAVAKIVFAVGFMIYVHSGEELTFLQVGMVETELAREDDIPITFVVL		367
QY	352	SK 353		
DB	368	PK 369		

```

RESULT 2
hypoetical protein At2g14060 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C:Accession: F84513
C:Gene: T23K23.11
R:Info: X: Kaul, S.; Rounsley, S.D.; Shea, T.P.; Bentol, M.L.; Town, C.D.; Fujii, C.Y.;
M.: Koo, H.; Molitor, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;
Cuss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Ventner, J.
Nature 402, 761-768, 1999
A>Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A444201 MIMD-20083487
A:Accession: F84513
A>Status: Preliminary
A:Molecule type: DNA
A:Residues: 1359 <STD>
A:Cross-references: GB:AE002073, NID:q4388826, PIRN:AA019781.1, GSDB:G000139
C:Genetics:
A:Gene: At2g14060
A:Map position: 2

Query Match 30.98; Score 570; DB 2; Length 359;
Best Local Similarity 36.98; Fred. No. 3.4e-38,
Matches 136; Conservative 64; Mismatches 137; Indels 32; Gaps 10;

UY 2 MNRGESSYVQNSSTFOGVASMAQF-ALDNAETLFSRDFHUALNAALGCAAGPNTF 60
DB 1 MKGGTGHSTVATNSHYQSRVYETQPLVIEVREMLKNGF-CGCKVADLGGSTGONTV 59
UY 61 AVIST/KMMKKKCPENLQTELEQVYINDP FGNDFNTLKK-----GLSEVIGNKE 113
DB 60 LAMSAIATVTMESYQSKNPEIDCYLNDPNDPNTTKLFHSFOEKLKPEVKK--- 116
UY 114 FVPCVWCVGSPGRLFPKNSLHVHSSYVHWLTQAPKCLTSRECLANKKITYSKT 173
DB 117 ---WFVSGVPSFSYSLRFPKRSLLFVHSAPSIHMLSKIPGLES-----NKSIIKYP 167
UY 174 SPVVRKAVYLSQFHEDETFMELNARSQEVNPGFVMTLLPGKQSPDSMOSCF-TWELLA 232
DB 168 YPSNVYKSYLNGKFIPESLFKMPSSEVNHGMVLTFCQPKVSDLS-KDCQVMSLIS 226
UY 233 MAIAEISQGLIDEDKLDFTNINISYFASLEVDYKRTSTFTTHESGT-----LIS 286
DB 227 DCLDLIDSEGVDSQKSMKSEFMPYNNEEVEEFILKESFEITIKEDHVPYKIDR 286
UY 287 VEMQEND-KWVRGEKFTKVRATFEPPIISNOGPEIMOKLYOKFTHIVSDLE-AKLPK 343
DB 287 FEEDHDSIQLEAFIKHASMAGCTTFPLVAHHCDAITFPVPRVYAHYAKYISVENHR 346
UY 344 TTSIIILVLS 352
DB 347 NMILIVVVS 355

RESULT 3
hypoetical protein T23K23.11 [imported] Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: F96703
C:Gene: T23K23.11
R:Info: X: Ecker, J.P.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso-
Chen, C.W.; Chung, M.K.; Conway, A.B.; Conway, A.R.; Greasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Hulzar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Phaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Liu, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Matli, P.; Martzali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Ventner, J.C.; Davis, R.W.
A>Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MIMD:21016719

```

```

A:Accession: F96703
A>Status: Preliminary
A:Molecule type: DNA
A:Residues: 1363 <STD>
A:Cross-references: GB:AE005174, NID:q6553892, PIRN:AA016558.1, GSDB:G000141
C:Genetics:
A:Gene: T23K23.11
A:Map position: 1

Query Match 27.48; Score 524.5; DB 2; Length 363;
Best Local Similarity 33.48; Fred. No. 6.3e-33;
Matches 115; Conservative 73; Mismatches 120; Indels 47; Gaps 8;

UY 2 MNRGESSYVQNSSTFOGVASMAQFALDNAV-----ETLSRDFHUALNAALGCA 55
DB 13 MSGGDPNSYSKNSHIQKRTSLIKEDLDVLEKNAKLISSDSN--TFRIADGAT 70
UY 56 GPNFAVISTIKMMKKKCPENLQTELEQVYINDP FGNDFNTLFPKLSSEVIGNKEFV 115
DB 71 GPNTEFLVDNLIKSIETSLKSNSSKPELVFNNDLPNDPNTLFTSLP-----GDR 122
UY 116 FCIWVSPGSEFHGRLFPKNSLHVHSSYVHWLTQAFGEISFEDLANKKITYSKTSP 175
DB 123 SYLWVGVGSEYGVHLDPQSSVHIVIMGAVHMLSSVEKVEYLSKSKAMKCKVHVSMAAD 182
UY 176 PVVPRVATISQFHPFMTPLNARSQEVNPGFVMTLLPGKQSP-SDMSQFTWELLAMA 234
DB 183 FVVR-ATVPCGPRMFKETLRFAPATEIVSGLLVWGMGSLPKMPSNLADSIYVSMATV 241
UY 235 IAEIVSQGLIDEDKLDFTNINISYFASLEVDYKRTSTFTTHDIEGFIHDSVEMQNDK 294
DB 242 LTQHSSELLISEQVDFTNIPVATPEEYTVLVVNGCCTV-----EEMELMDPTA 293
UY 295 WVRGEKFTK-----VKATFEPPIISNOGPEIMOKLYOKFT 330
DB 294 WIK--PPTNEDVHMVVCIKATMGSLFINHGHDLDVDFRLT 336

RESULT 4
hypoetical protein T23K23.11 [imported] Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: A86285
C:Gene: T23K23.11
R:Info: X: Ecker, J.P.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso-
Chen, C.W.; Chung, M.K.; Conway, A.B.; Conway, A.R.; Greasy, T.H.; Dewar,
ansen, N.F.; Hughes, B.; Hulzar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Phaykin, E.; Kim,
C.A.; Li, J.H.; Li, Y.; Liu, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Matli, P.; Martzali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Ventner, J.C.; Davis, R.W.
A>Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MIMD:21016719
A:Accession: A86285
A>Status: Preliminary
A:Molecule type: DNA
A:Residues: 1351 <STD>
A:Cross-references: GB:AE005172, NID:q6104411, PIRN:AA019781.1, GSDB:G000141
C:Genetics:
A:Gene: F9L1.6
A:Map position: 1

Query Match 27.28; Score 499.5; DB 2; Length 351;
Best Local Similarity 33.68; Fred. No. 7.5e-32;
Matches 122; Conservative 73; Mismatches 141; Indels 27; Gaps 10;

UY 2 MNRGESSYVQNSSTFOGVASMAQFALDNAVETLFSRDFHUALNAALGCAAGPNT 59
DB 1 MNRGESSYVQNSSTFOGVASMAQFALDNAVETLFSRDFHUALNAALGCAAGPNT 59

```







G64309  
 hypothetical protein homolog M10079 - Methanococcus jannaschii  
 CSpecies: Methanococcus jannaschii  
 CDate: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 21-Jul-2000  
 CAccession: G64309  
 R:Rault, C.J., Miller, O., Olsen, G.G., Zhou, L., Fleischmann, R.D., Sutton, G.G., Blake, R., Reich, C.T., Ouellet, R., Kikinis, E.F., Weinstock, K.S., Merrick, J.M., Chodura, A., Rison, J.D., Sadow, P.W., Hanna, M.C., Cotton, M.D., Roberts, K.M., Holsclaw, M.A., Science 273, 1058-1073, 1996  
 A:Authors: Kaine, R.P.; Petrovsky, M.; Kien, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C.A.; Miller, C.; Olsen, G.G. The complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii  
 A:Reference number: G64309; M010-96337999  
 A:Accession: G64309  
 A>Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Position: 1-360 cpm  
 A:Cross reference: GI 067463 (L77117) R10350001, E100 AAB08094, E10350000  
 C:Genetics:  
 A:Map position: FEV73806-72664  
 A:Start codon: TTG

Query Match 57% Score 106, 108.2, Length 120,  
 Best Local Similarity 20.1% Pred No. 0.78,  
 Matches 64, Conservative 66, Mismatches 127, Indels 62, Gaps 14,  
 Oy 63 ISTIKMKKKCEELNGCL-----ELQVNLPLSGNEVTLFKLSSEVIGNKEVEVY 118  
 Db 1 MSEIMKELKIPERINSYFEPFEDIALSTLANE-HIVF----- 41  
 Oy 119 VMGVPSPGSHGRIFPNSIHVHSSVHMGTQAPKITSPECL-----ALNKGKIYS 171  
 Db 42 -IGNPGVANSQILRALASH-INANVFEKLITR-----FTTEDELPLPLSKELKNDPRVR 95  
 Oy 172 KTF---PVHEAVLSQHECTFPLNAR---SGHVINGQVW-----LILKPGQSD 218  
 Db 96 KISGPTAEIAFLDEVFANSILNALISINERLYHNGDLKRVPLISLPGASNELPE 155  
 Oy 219 PSMQSGPTMELLAMIAELVS-QGLDEKLEFNISYFSLSEVNDIVERGSEFTID 277  
 Db 156 ENELAFYDFELFRKVPKIPSCENLVKIKIDFEYKPTTISIELPKMKRANEVDIF 215  
 Oy 278 HTEFPL SEVNGEMKVEKPELVVAF-TEPISNGGFEIMKLYDKFTIHVS 335  
 Db 216 NITGVVDIKKKKSNHITISDPFRKSKAIKCAVINGFEAFITF-LETLFIHFW 273  
 Oy 336 ELEAKLPKTSILVLSKI 354  
 Db 274 DIDP-----LLIVSKV 284

## RESULT 11

T41720  
 hypothetical protein SPAC821.09 - fission yeast (Schizosaccharomyces pombe)  
 CSpecies: Schizosaccharomyces pombe  
 CDate: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 03-Dec-1999  
 CAccession: T41720  
 R:Ragot, M., Lyne, M., Rajandream, M.A., Barrell, B.G.  
 A:Submitted to the EMBL Data Library, October 1999  
 A:Reference number: Z22612  
 A:Accession: T41720  
 A>Status: preliminary; translated from cDNA/EMBL/GenBank  
 A:Molecule type: DNA  
 A:Residues: 1-1016 cpm  
 A:Cross-reference: EMBL AL12770, F10N YAR57443.1, GSPFB GNG006, SPH:STAF421.09  
 A:Experimental source: strain 972h; cosmid 7b21  
 C:Genetics:  
 A:Gene: SPH:SPAC821.09  
 A:Map position: 1

Query Match 55% Score 101.5, 108.2, Length 1016,  
 Best Local Similarity 18.4% Pred. No. 5.6;

Matches 70, Conservative 43, Mismatches 127, Indels 139, Gaps 14;  
 Oy 9 SSYANSSPPEVYASMAQ--PALENAVEFLSP-----DPHQALNMAAD-L 51  
 Db 2 SSYASPEFPALEHLSIAQSPITKPKFKSTGNSISIKKKQNVFVSVIDINPASYE 61  
 Oy 52 GTAAGNIAVAVISIKMKKKKSLNVLHLEVYINLNGNINLPKQ LSS 105  
 Db 62 GIVSHVILVAVSI-----LSSLELNKLEEDNLSSLEPNIWALVYVWSS 112  
 Oy 106 EVLVNRC-----EEVLYVWVGSPHGLPFPNSIHVHSSVHMGTQA 151  
 Db 113 DPMGSGIANTHNGVMDPTIPSY-----YE 142  
 Oy 152 PKQTSPEALNKKKTYSKTSPPVPPATVSGHEPTMFAANSQVPPKQVYIL 211  
 Db 143 PHLVS LSSVSSLSNHFIVE IYMSAPET----- 175  
 Oy 212 KQWSSSLSUMSGEWEELAMIAELVSQGLDEKLEFNHISYASLEKLYEPD 271  
 Db 176 -----PSWSSSMQLT-LTEGMAYTAYTNATPQISSILYINDVEVP 219  
 Oy 272 GPTTHHEFTTHQSEVEMENRW--VPGKPTKVVAFTPEPISNOF----- 317  
 Db 220 QISAVAKYR VFMNHWMLLYVQSLLEISLKWSSGTFNYQIAKITL 273  
 Oy 318 GPELMKLUKFTHIYVSD 337  
 Db 274 GCEALVLYTVAGVITGI 293

## RESULT 12

T15137  
 hypothetical protein T28F2.3 - Caenorhabditis elegans  
 CSpecies: Caenorhabditis elegans  
 CDate: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 20-Sep-1999  
 CAccession: T15137  
 R:Madson, C.; Promick, B.  
 A:Submitted to the EMBL Data Library, April 1997  
 A:Description: The sequence of C. elegans csm4 T28F2.  
 A:Reference number: Z18300  
 A:Accession: T15137  
 A>Status: preliminary; translated from cDNA/EMBL/GenBank  
 A:Molecule type: DNA  
 A:Residues: 1-319 cpm  
 A:Cross-reference: EMBL AF000196, M10-3204741, F10-3204754K, F10N AAB53554.1, GSPFB  
 A:Experimental source: strain Bristol NC3; clone T28F2  
 C:Genetics:  
 A:Gene: CSDP:T28F2.3  
 A:Map position: 1  
 A:Inserts: 90/2; 189/3; 266/3; 295/3

Query Match 54% Score 99, DB 2; Length 319;  
 Best Local Similarity 20.3% Pred. No. 1.9;  
 Matches 75, Conservative 56, Mismatches 106, Indels 132, Gaps 18,

Oy 1 FMNVESESSYVANSFTQVYVMAQPA-----LENAVETLSR---DFHL 43  
 Db 3 FLSGLDQ--GFAQ-----LKQKALMAFSLAELNQLKEMKRESAFTGINKKLLSLG 56  
 Oy 44 GAINAAIAQAAAP-NFAVAVISIKMKKKKPELNVLPIQVYINLPGNINLPKQ 102  
 Db 57 KLVNPSLSTGSPIDIVPITAFGEHLQNAHFVYESTGEKAVND--GNSIIMREG 114  
 Oy 103 LSSVAVGNKREVPVVMVYVNSPFGKLPFN-----SIHVVSSVHMT 144  
 Db 115 NSSRLATSPPEQYHIDAV--NFMATPMMGSEHITIGVGVAGEMHIDRN----- 165  
 Oy 150 GAPPALSPPELALNKKIYISKTSPPVVEAVLSQFED---FTMHNAVSQVEVPMQ 206  
 Db 166 --TPPAIMAIKQPNVIAIA-----VFINPSHQAANVFSPILNLPQ----- 207

[illegible][illegible]

```

Query March 1997 Score 307.16 E-01 Length 1706
Best Local Similarity 21.381 Prod. No. 261
Matches 77; Conservative 54; Mismatches 129; Indels 102; Gaps 17.

QY 28 ALENAAV--ETLFSSRD--PHIQAINAALDSCAA-----GNTRAV-----IS 64
      ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 68 ALDEIESEALLSDIFQINLEEAVDLLIAGEAOKIHEFGINKHLLAVVGYADAKILIA 127
      ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 65 IINPMKKRIKELNQLLEELGYVNLINLEND--PHILPKNQSSAVYNNKTHVNC--YVKA 121
      ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 128 IIRKMLKWLQKEDL---PQPLSKELNLEINELVINKLVQGLINLVQSEFRKQSPDNG 184
      ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 122 VPQSGHGRLEPRNSLIHLVHSY--SVHWITQAFKGLTSREIALNKKRIYISKTSPP--- 176
      ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 185 IEGSKHAKIVKNAITFISVYIGLSMIGPDPG---NAVISINHPQIVAAVPPKKI 249
      ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 177 -----VNEATVSCHEFTPTPTFLNARQGVNPPKQWLLIKSGVCSNDSDMG 224
      ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 240 SCSSITAMISLAVTTSFVSLQVND-----AQGVLTIN---KTAHINKTIDMSVGA 286
      ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 294 SFTPTWETIMAAIATPVSGELTETPTETETATSYGAATVQKQVYVQVETITQVPTG 282
      ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 287 MCGTGLQACAVALKRSTASS-----PSHQALETENIKVDYK----- 421
      ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 292 IMLVENVQLMEKVFQJDKITVVAATLILINRQETELKGLTILITVNSALLATG 344
      ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 422 -----VTPGSPINMAATVETFGALITKSTIRFVADHQMIDELLKQIISVPAAKIM 470
      ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 343 KT 344
      ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 371 ET 372
      ||| : : : : : : : : : : : : : : : : : : : : : : : : : :

```

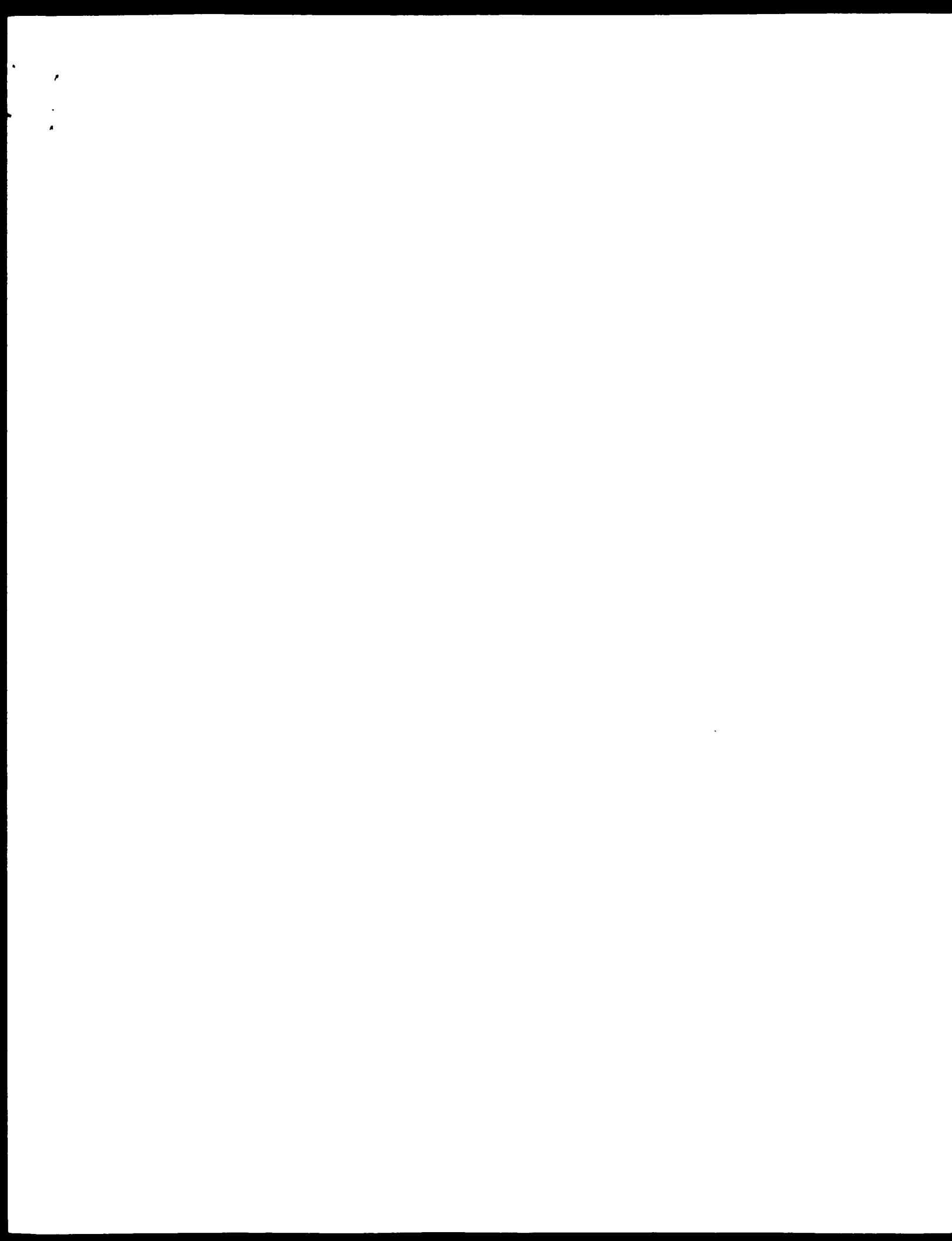
RESULT 15  
S44194  
poliprotein - echovirus 12  
C:Species: echovirus 12  
C:Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 26-Aug-1999  
C:Accession: S44194  
R:Kraus, W.; Nelsen-Salz, B.E.  
submitted to the EMBL data library, February 1994  
A:Reference number: S44194  
A:Accession: S44194  
A:Status: preliminary  
A:Molecule type: genomic RNA  
A:Residues: 1-2193 <KRA>  
A:Cross\_references: EMBL X77708, NIT: q2058304, FIDN:GAA54763.1, FID: q474422  
C:Superfamily: poliovirus genome poliprotein  
C:Keywords: genome linked protein, phosphoprotein, poliprotein  
F152Y/Binding site: phosphoryl RNA (Tyr) (covalent) #status predicted

```

OY 29 LENVETLESFDPHIGALNADIGCAAGPNTAV-----ISTKEMMEK-- 72
Db 182A LEPYAV-----YVTRGATFATIGTISAGYVAIGIKKRIIISKRIKDI IKIKPMEDKYG 1880
OY 73 -----KCRELNCOTLEQVYLNDLFCNDENTLFKG---LSSFY 107
Db 1881 LNLPMVTYVDELPSSPKVAKGKSPITEASSINDSVAMPQTPSNYKTFHINPGITVQSA 1940
OY 108 LGNKCFPVPCYVMGVPSSEPHGLPEPNSLHLVHSSYSVHMLTQAP-----KGL 155
Db 1941 VG--CD-----PDLFWSKIPVMDGHLIAPFVSGYDASISPVWFACTIKLILEKLOY 1989
OY 156 TSPRGLAINCKRTIYISKTSPPVVPFVATISQPHEDF-----TMFLNAPS 198
Db 1990 THRF-----TWYID-----YLCNSHLVPPDKHYFVVGMPGSGSSTISFNSKI 2032
OY 199 QEVVPNGWVLLIIG-----PQSDPSIMQSGFTWELLAMAIJEL-VSQGLI--DEKRL 249
Db 2033 NNLIIPTIMKVVYKGTITIGFPMIATGTOVIASTPMTDASIIATFACPGYGIIMTPADKG 2092
OY 250 DTFN-----IPSYFASLEF-----VKDIVE-----PDGSPFTIHIEGFPYD 285
Db 2093 FPFNEVITWIVTEFLKPYFPADEQYPLVHPVMPMKDIHESIPTKDPKNTQDHVPSLICIL 2152
OY 286 SVE--MGENCKWVCGEFTFYVVPATIEPLISNCGSEIMMKLYDKF 329
Db 2153 AMHNGEDVFPFPIPKIIPSVVVGPCLTLPABST-----LPPKWLUSP 2193

```

Search completed: August 31, 2001, 15:01:17  
 Job time: 2130 sec



GenBank version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd

CM protein - protein search, using SW mode!

Run on August 31, 2001, 14:34:27; Search time: 08 seconds  
(without alignments)  
281,065 Million cell updates/sec

Title: US-09-577-657A-1

Perfect score: 1847

Sequence: 1 FMNRTGSSSYAQMSTPTQ

LEAKIKRTSLIIVLSKILW; 456

Scoring table:

BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 197339 seqs, 20590346 residues

Total number of hits satisfying chosen parameters: 197339

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: Issued Patents AA: \*  
1: /cgn2\_6/prodata/2/1aa/TA.COMB.pep.\*  
2: /cgn2\_6/prodata/2/1aa/5A.COMB.pep.\*  
3: /cgn2\_6/prodata/2/1aa/6A.COMB.pep.\*  
4: /cgn2\_6/prodata/2/1aa/7B.COMB.pep.\*  
5: /cgn2\_6/prodata/2/1aa/PTCS.COMB.pep.\*  
6: /cgn2\_6/prodata/2/1aa/hb-hfltest.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No	Score	Query Match	Length	DB ID	Description
1	97.5	53	2763	US-08-496-944-2	Sequence 2, Appl
2	85	46	885	US-08-484-105-14	Sequence 14, Appl
3	85	46	885	US-08-484-106-14	Sequence 14, Appl
4	81	44	553	US-09-068-655-6	Sequence 5, Appl
5	80	43	441	US-08-114-309A-4	Sequence 4, Appl
6	79.5	43	118	US-08-545-809A-145	Sequence 145, App
7	79.5	43	2465	US-08-596-291-3	Sequence 3, Appl
8	79.5	43	2465	US-09-100-804-3	Sequence 3, Appl
9	79	43	428	US-07-882-790-4	Sequence 4, Appl
10	79	43	1429	US-08-540-804-14	Sequence 14, Appl
11	79	43	1429	US-08-118-245-14	Sequence 14, Appl
12	79	43	1429	US-08-521-872-14	Sequence 14, Appl
13	79	43	1420	US-08-590-399-14	Sequence 14, Appl
14	78.5	43	2296	US-08-286-819A-27	Sequence 27, Appl
15	78.5	43	2296	US-08-286-819A-27	Sequence 27, Appl
16	78	42	438	US-08-577-492-38	Sequence 38, Appl
17	78	42	968	US-09-180-439-3	Sequence 3, Appl
18	78	42	968	US-09-180-439-3	Sequence 3, Appl
19	78	42	1016	US-09-180-439-8	Sequence 8, Appl
20	78	42	2296	US-08-480-604A-13	Sequence 10, Appl
21	77.5	42	2366	US-08-405-490A-18	Sequence 10, Appl
22	77.5	42	199	US-08-663-566A-9	Sequence 9, Appl
23	77.5	42	501	US-08-023-610-9	Sequence 9, Appl
24	77.5	42	501	US-08-288-062A-9	Sequence 9, Appl
25	77.5	42	501	US-08-288-062A-9	Sequence 9, Appl
26	77.5	42	501	US-08-288-062A-9	Sequence 9, Appl
27	77.5	42	501	US-08-804-372A-7	Sequence 7, Appl

# ALIGNMENTS

RESULT 1  
US-08-496-944-2  
Sequence 2, Application US/08496944  
Patent No. 6040496  
GENERAL INFORMATION:  
APPLICANT: Law, Marcus D  
TITLE OF INVENTION: Use of translationally altered RNA to  
TITLE OF INVENTION: Control resistance to Mammalian Virus and Other  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CIBA-Geigy Corporation  
STREET: 7 Skyline Drive  
CITY: Hawthorne  
STATE: NY  
COUNTRY: USA  
ZIP: 10532  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-TS/MS-DOS  
SOFTWARE: PatentIn release #1.0, Version #1.308  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/496,944  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY, FIRM, INFORMATION:  
NAME: Elmer James Scott  
REGISTRATION NUMBER: 36,129  
REFERENCE/DOCKET NUMBER: CGC 1814  
INFORMATION FOR SEQ ID NO.: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2763 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULAR TYPE: protein  
US-08-496-944-2

Query Match: 534; Score 97.5; DB 3; Length 2763;  
Best Local Similarity 19.88; Pred. No. 1.3; Mismatches 93, Gaps 17,  
Matches 745; Conservative 58, Mismatches 144, Indels 93, Gaps 17,  
5: GCGSSYVAGSEFFVYASMAQALERAVE-ITFSEFHILGATNMAADGACAHIFAVI 63  
147: GYKSKIMNAAFIKDSKAFVYIGHLYHLHVAWG ..... FVV 203  
67: 64 STIRMEKRC... FENCGTLELVYINIEFGDFNTFKGLSS-----EVINKKEE 114

```

Db 2014 NLLKNGMOQCVVYVDEEFIRSLINAAVSAITYGKKRVFENFESSEKELIEMESCEP 2073
QY 115 VPCVMEV-PGSHGRLPFRNSLH-----VHSYSVH-----146
Db 2074 IYXKXGWNKSLKAEIRPLEKTMANKTRIFVAPLETLGGVYCVDEENNQYSHHLEP 2133
QY 147 -MLTQARK-GLTREGIALNKRY-----ISKTPPVVREAYLS--QFHEPF--190
Db 2134 PWVGTIRFYGWMLLEKLEPGCWVYCDAGSQPDSLTFYLLNAVNLRLQFHEMDIG 2193
QY 191 -TMEIANSQIV---VENCWVIIRPGYCSIPDSMOSCFWELLAMIAELVSGCL 244
Db 2194 AQMKNYTELIVYPIAPDSIVYKFRKGNNSQSPSTVYDNLVIIFENYA-MLSSG-1 2251
QY 245 DEDKLPFNIHSYASLEEKDIVERGSPFIIDHIE-----GFLDSVEMQENDK- 295
Db 2252 KEHLD-RCGNMFRANDDLIAVHIFERITLDFQNHGNLNFEFSIKKQKSLMF 2509
QY 296 --VQGEK 301
Db 2310 MSTGQIKY 2417

```

```

RESULT 2
US-08-484-105-14
: Sequence 14, Application US/08484105
: Patent No. 5589341
: GENERAL INFORMATION:
: APPLICANT: STILLMAN, Bruce
: APPLICANT: BELL, Stephen P
: APPLICANT: KOBAYASHI, Ryuji
: APPLICANT: RINE, Jasper
: APPLICANT: FOSS, Margit
: APPLICANT: MCNALLY, Francis J
: APPLICANT: LAURENSEN, Patricia
: APPLICANT: HERSKOWITZ, Ira
: APPLICANT: LI, Joachim J
: TITLE OF INVENTION: ORIGIN OF REPLICATION COMPLEX GENES
: NUMBER OF SEQUENCES: 24
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: FLEHR, HOBACH, TEST, ALBRITTON & HERBERT
: STREET: 4 Embarcadero Center, Suite 3400
: CITY: San Francisco
: STATE: California
: COUNTRY: USA
: ZIP: 94111-4187
: COMPUTER READABLE FORM:
: MEDIUM TYPE: floppy disk
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/484,105
: FILING DATE:
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Osman Ph.D., Richard Aron
: REGISTRATION NUMBER: 36,627
: REFERENCE/DOCKET NUMBER: A-59032/DTF/PAO
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (415) 494-8700
: TELEFAX: (415) 494-8771
: TELE: 910 277299
: INFORMATION FOR SEQ ID NO: 14:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 885 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-08-484-105-14

```

Query Match 4.68; Score 85; DB 1; Length 885;  
Best Local Similarity 19.08; Pred. No. 4.4;  
Matches 60; Conservative 50; Mismatches 95; Indels 110; Gaps 10;

```

QY 50 DIFGAAGNTFAVIST--IKFMKEKCPETNFOTLEIYVINDLPG-----NDENTL 99
Db 391 DIDLIAALFFRRFYTSAGKMKETTFSKYKGLNSFNKKEEIVKAVPDNLLPAPENEFASI 450
QY 100 FKGLSSFVIGNKCEVPQYMGVPG-----SPHGLIFPR-----NSLH 137
Db 451 YLISYSAIEAGISTSI--YIAGTPGVGKITLVREVKDLMISADQKELRPQYIEINGLK 508
QY 138 LVHS SYSHWULTQAPKGLTSFPGIAIANGKTYIKTSPVYVREAYLSGFHECTMFLN 195
Db 509 IVKASUSTEYFMOKISGKILIS--GAAMESLFEYFNKV--PAIKKRPVIVLIDELDA-IV 563
QY 196 APSQVYVNPYMWLILPGPGCSPPDSMOSCFW-----ELAMAIAYELVSGCLIDEPKL 249
Db 564 SKSQDV-----MYNFFNMATYSNAKLIVAVA-----590
QY 250 DTFNIPSY--FASL-----HVKQIVF--KQSTTIAHISF--282
Db 591 NTLDPERHLGNKLSRIGFTFRIMFTGYTHEELRTIINLRLMYLNSSFYVDPFGSSYM 650
QY 283 --GLDSVEMQENDK 294
Db 651 ISPSSTIIFDEDERK 665

```

```

RESULT 3
US-08-484-106-14
: Sequence 14, Application US/08484106
: Patent No. 5614618
: GENERAL INFORMATION:
: APPLICANT: STILLMAN, Bruce
: APPLICANT: BELL, Stephen P
: APPLICANT: KOBAYASHI, Ryuji
: APPLICANT: RINE, Jasper
: APPLICANT: FOSS, Margit
: APPLICANT: MCNALLY, Francis J
: APPLICANT: LAURENSEN, Patricia
: APPLICANT: HERSKOWITZ, Ira
: APPLICANT: LI, Joachim J
: APPLICANT: GAVIN, Kimberly
: TITLE OF INVENTION: ORIGIN OF REPLICATION COMPLEX GENES
: NUMBER OF SEQUENCES: 24
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: FLEHR, HOBACH, TEST, ALBRITTON & HERBERT
: STREET: 4 Embarcadero Center, Suite 3400
: CITY: San Francisco
: STATE: California
: COUNTRY: USA
: ZIP: 94111-4187
: COMPUTER READABLE FORM:
: MEDIUM TYPE: floppy disk
: OPERATING SYSTEM: IBM PC compatible
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/484,106
: FILING DATE:
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Osman Ph.D., Richard Aron
: REGISTRATION NUMBER: 36,627
: REFERENCE/DOCKET NUMBER: A-59032/DJB/PAO
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (415) 494-8700
: TELEFAX: (415) 494-8771
: TELE: 910 277299
: INFORMATION FOR SEQ ID NO: 14:
: SEQUENCE CHARACTERISTICS:

```

LENGTH: 885 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-484-106-14

Query Match 4.6%; Score 85; DB 1; Length 885;  
Best Local Similarity 19.0%; Pred. No. 4.4;  
Matches 60; Conservative 50; Mismatches 95; Indels 110; Gaps 16;

QY 50 LUUCAGUNITAVIST---IKKMKKKKRLNQLTLELYLNDLGC-----NDFNRL 99  
DB 391 DLDAIALEFRPRTVSAKGMETTFISKVKKOLNSRNSKEETVKAADDNVLPAENEFASI 450  
QY 100 FKGLSEVYGNKCEVPDQYVWGVPG-----SPHGRLFP-----NSLH 137  
DB 451 YLSLYSAIEAGTSTST---YIAGTGVGKTLTVREYVADMTSMADQKELPRFYELNGLK 508  
QY 138 LVHS--SYVHWLTQARKGLTSPGTLANKGITYISTSPVYVFAVLSQFHELPIMFLN 195  
DB 509 IVKASDSEYEVWQKTSCKRLTS--GAAMESLEFYFNKV--PATKKRPVVLIFELNA-LV 563  
QY 196 ARSQEVYVNGCMVJLNGRQCSQPSQMSQCFW-----ELLAMATAELVSGQLIEDKL 249  
DB 564 SKSDVY-----MYNEFNMATYSMAKLIVAVAA----- 590  
QY 250 DTFNIPSY-----FASL-----EYKADIVE-----RDQSPTIDIEGF--- 282  
DB 591 NTLDLPEFHLCNKTISSRIQFTRIMFTGYTHRLRTIINLEPLKYTLNSSFVVDPEFTSSVM 650  
QY 283 ---DIDSVEFMQPNK 294  
DB 651 ISPOSTIETDEEK 665

RESULT 4  
US-09-068-655-6  
Sequence 6, Application US/09068655A  
Patent No. 6136579  
GENERAL INFORMATION:  
APPLICANT: JACKSON, Ronald James  
TITLE OF INVENTION: METHOD OF PRODUCING ALPHA 2,4-SIALYLTRANSFERASE  
FILE REFERENCE: 026579-186  
CURRENT APPLICATION NUMBER: US/09/068,655A  
CURRENT FILING DATE: 1998-08-31  
EARLIER APPLICATION NUMBER: AU PN 6587  
EARLIER FILING DATE: 1995-11-15  
EARLIER APPLICATION NUMBER: PCT/AU96/00725  
EARLIER FILING DATE: 1996-11-15  
NUMBER OF SEQ ID NOS: 13  
SOFTWARE: Patent In Ver. 2.0  
SEQ ID NO 6  
LENGTH: 553  
TYPE: PPT  
ORGANISM: Myxoma virus  
US-09-068-655-6

Query Match 4.4%; Score 81; DB 4; Length 553;  
Best Local Similarity 20.6%; Pred. No. 5.6;  
Matches 65; Conservative 40; Mismatches 97; Indels 114; Gaps 13;

QY 62 VISITKMKKKKRLNQLTLELYLNDLFGNDPNTLFGSLSEVY-----GNKCEVP 116  
DB 144 VTTIHTIVSVLRQPNKRIELCLNLTSSDPLNLTEDVCAVVLWLKNNMELDC 203  
QY 117 CYV-----MGVFGSHGKLFKRNLSLHVSYSVHWLTQARKGLSKGLSLNGL 166  
DB 204 SVLEQVMSLSMSVKNLILKTCTFNKRYV-----QSLAKLIDHSPPPPTGTTISGGR 259  
QY 167 KIVISKISPV-----VPEAVI-----SDFHR-----DFIMFLNKKSEVVP----- 203

DB 260 KYDDITSSPELVKSPVDVWTVTSYLPTRQPFVAVLDEVVVVGGLQDSVSVSS 319  
QY 294 -----NQMVLIR-----GPGCSQPSQMSQCFWELLAMALA 236  
DB 320 YIVKTNKKKEPLKSRHRHSGLVVILKLIIVSGSKN-----SYLKIVVYVW 568  
QY 237 ELVSQHTIIRIKLIPFNIPSY-----FASLEVKD-----IVEKGSFPIIDHIEGFO---- 283  
DB 369 -----PIYATWRKLSIREAFITVGAAYVFNKVIYIGISVSEPSR 410  
QY 284 ---LDSVEMQENDKVV 296  
DB 411 LFTDLYVFTQNNKVV 426

RESULT 5  
US-08-314-309A-4  
Sequence 4, Application US/08314309A  
Patent No. 5677141  
GENERAL INFORMATION:  
APPLICANT: TSUGAI, TAKAO  
APPLICANT: FUKAGAWA, MASAO  
APPLICANT: IWAMI, MORITA  
APPLICANT: ARAMORI, ICHIRO  
APPLICANT: KUDO, HITOSHI  
TITLE OF INVENTION: PROCESS FOR PRODUCING 7-AMINOCEPHEM  
TITLE OF INVENTION: COMPOUND OF SALT THEREOF  
NUMBER OF SEQUENCES: 34  
CORRESPONDENCE ADDRESS:  
ADDRESSER: ORION, SPIVAK, MCLELLAND, MATHER & NEUSTADT,  
STREET: 1755 S. Jefferson Davis Highway, Suite 400  
CITY: Arlington  
STATE: Virginia  
COUNTRY: U S A  
ZIP: 22202  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/314,309A  
FILING DATE: 30-SEP-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/631,906  
FILING DATE: 21-DEC-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: O'Brien, NO. 5677141man F.  
PRTSTATION NUMBER: 24,618  
REFERENCE/WORKER NUMBER: 18-863-0 CONT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 413-3000  
TELEFAX: (703) 413-2220  
TELEX: 248855 OPAT DR  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 441 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-314-309A-4

Query Match 4.3%; Score 80; DB 1; Length 441;  
Best Local Similarity 22.0%; Pred. No. 4.9;  
Matches 35; Conservative 30; Mismatches 54; Indels 40; Gaps 7;  
QY 139 VHSYSVHWLTQARKGLTSPGTLANKGLIYISTSPVYVFAVLSQFHELPIMFLNARS 198  
DB 25 INQPSQGLHTNANMVLIR-IAVILGALPAALAAV-----TLCK 64

QY 199 GVAVNVMVITIGGQSPRSGMGTFTW--ELLAMAIELVSGIIECKLDTNITS 256  
 DB 65 REVVNRKIVTLKGG-----ASNDSHISVNSDIHKRSLSRSTAGIEKFFHIDTEN--A 117  
 QY 257 YFASLEEKDIVERGCTTIDHIEGFLDSVEMOEHEK 295  
 DB 118 YVGEFET-----TIEIKN-NDVLEVEEDIM 145

RESULT 6  
 US-08-545-809A-145

Sequence 145, Application US/08545809A

Patent No. 6096878

GENERAL INFORMATION:

APPLICANT: Honjo, Tasuku

APPLICANT: Matsuda, Fumihiko

TITLE OF INVENTION: HUMAN IMMUNOGLOBULIN VH GENE

NUMBER OF SEQUENCES: 145

CORRESPONDENCE ADDRESS:

ADDRESSEE: Fish & Richardson, P.C.

STREET: 225 Franklin Street

CITY: Boston

STATE: MA

COUNTRY: US

ZIP: 02110-2834

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: Windows95

SOFTWARE: FASTSEQ for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/545,809A

FILING DATE: 27-MAR-1996

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/JP93/00003

FILING DATE: 10 MAY 1993

ATTORNEY/AGENT INFORMATION:

NAME: Freeman, John W.

REGISTRATION NUMBER: 29,066

REFERENCE/WORK NUMBER: 08501/004001

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617-542-5070

TELEFAX: 617-542-8906

TELEX: 300154

INFORMATION FOR SEQ. ID NO: 145:

SEQUENCE CHARACTERISTICS:

LENGTH: 118 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-545-809A-145

Query Match 4.38; Score 79.5; DB 3; Length 118;  
 Best Local Similarity 26.18; Pctd. No. 0.69;  
 Matches 31; Conservative 21; Mismatches 40; Indels 27; Gaps 5;

QY 99 LFKGISSEVIGNKCEEPVYVMGPGSFHGRLEPRNSILH-----VSSSYVHLTQA 151  
 DB 14 LFKVVOCEV-----OLVESGGGLVPGGSLRLSCAASGFTSSYAMHWROA 60  
 QY 152 P-KGLTSHREGIALNKCKIYISKTSPPVVRAYLSQFHEDFTMPL--NAKSQEVVYNGC 206  
 DB 61 PKGLTEVVAISNSNGSTYYANS--VKGRFTISRDNSKNTLTLQNGSI,PAEFMAYVYC 116

RESULT 7

US-08-596-291-3

Sequence 3, Application US/08596291

Patent No. 5821075

GENERAL INFORMATION:

APPLICANT: GONZ, LEONEL JORGE  
 APPLICANT: SARAS, JAN  
 APPLICANT: CLAESON-WELSH, LENA  
 APPLICANT: HELDIN, CARL-HENRIK  
 TITLE OF INVENTION: PRIMARY STRUCTURE AND FUNCTIONAL  
 EXPRESSION OF NOVEL HIE SEQUENCES FOR NOVEL PROTEIN  
 TITLE OF INVENTION: TYROSINE PHOSPHATASES  
 NUMBER OF SEQUENCES: 4  
 CORRESPONDENCE ADDRESS:

ADDRESSEE: WOLF, GREENFIELD & SACKS, P.C.

STREET: 600 ATLANTIC AVENUE

CITY: BOSTON

STATE: MASSACHUSETTS

COUNTRY: USA

ZIP: 02210

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: P-C-P-S-M-S-1005

SOFTWARE: Patentlib Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/596,291

FILING DATE: 09-AUG-1996

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/115,573

FILING DATE: 01-SEP-1993

ATTORNEY/AGENT INFORMATION:

NAME: GATES, EDWARD R.

REGISTRATION NUMBER: 41,615

REFERENCE/WORK NUMBER: 10401/0000

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617/720-3500

TELEFAX: 617/720-2441

TELEX: 92-1742 ECKEIL

INFORMATION FOR SEQ. ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 2465 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US 08 596-291-3

Query Match 4.38; Score 79.5; DB 2; Length 2465;  
 Best Local Similarity 22.09; Pctd. No. 87;  
 Matches 46; Conservative 35; Mismatches 91; Indels 37; Gaps 8;

QY 2 MNRREBSYVAQNSSTFYVASMAGFALFNAYETLPSTQPHI,LAALNADL3AAGPNTFA 61  
 DB 966 MSKSYHDLISQASLYPRHKNVIVMDEPPQIVAEIVGKPSHOMSRSLASLACVTKLNNSK 1025  
 QY 62 VISTIKRMKKCKRELNCQLE--LOYLYNDLFGDNFN-----TLFK-----GL 103  
 DB 1026 SVASLNSPPEPKHESQSSSTENPPQAVYIVDIHKPWSIVSSPPPTTLVNI KPAKYGL 1065  
 QY 104 SSEYGNKCEE-----VPCYVMGVPQSEH5-----PLFPNSJHLVHSSY--SVHMLT 149  
 DB 1086 GFUJIGGEKMEIDLGIFISVAVGPAIDFHGCLKPKQDKLISVNSVSLSEVSHHAAIETLQ 1145  
 QY 150 GAKGLISPEGLANKKITYISKTS-SPV 177  
 DB 1146 NAPIDVT---LVISQPKKISKVPSTPV 1170

RESULT 8

US-09-100-804-3

Sequence 3, Application US/09100804

Patent No. 6066472

GENERAL INFORMATION:

APPLICANT: GONZ, LEONEL JORGE

APPLICANT: SARAS, JAN

APPLICANT: CLAESON WELSH, LENA



```

APPLICANT: HELDIN, CARL-HENRIK
TITLE OF INVENTION: PRIMARY STRUCTURE AND FUNCTIONAL
TITLE OF INVENTION: EXPRESSION OF NUCLEOTIDE SEQUENCES FOR N-WEEL PROTEIN
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESS: WOLF, GREENFIELD & SACKS, P.C.
STREET: 600 ATLANTIC AVENUE
CITY: BOSTON
STATE: MASSACHUSETTS
COUNTRY: USA
ZIP: 02210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patulin Polcage #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/100,804
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/596,291
FILING DATE: 09-AUG-1996
APPLICATION NUMBER: US 08/115,573
FILING DATE: 01-SEP-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/06943
FILING DATE: 01-SEP-1994
ATTORNEY/AGENT INFORMATION:
NAME: GATES, EDWARD R.
REGISTRATION NUMBER: 31,616
REFERENCE/WORKSHEET NUMBER: 10441/7003
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-720-3500
TELEFAX: 617-720-2441
TELEX:
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 2465 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-100-804-3

Query Match 4.38; Score 79.5; DB 3; Length 2465;
Best Local Similarity 23.08; Pred. No. 87;
Matches 46; Conservative 35; Mismatches 91; Indels 37; Gaps 8;

QY 2 MNGGEGSSVAQSSFTQVYASMAQPALENAVETLESKPHLCALMAALGCAQNPITFA 61
DB 966 MSSSYHRTSGASLYPIHKNTVIMPEPQVVAIVKPSHMSPSDAESLACVTKINNSK 1025
QY 62 VISTIKMKMKKRPINCOITF--IQVYINDLFGNDN-----TFK-----GI 103
DB 1026 SVASLINSRPERKHEDSSSTEDPGQAVYVDYHKRMSLVSPDEEITLVNKKDAKGL 1085
QY 104 SSEYIGKCE-----VPCYVWGVGSHG-----RFPNNSLHVHSSY--SVHMIT 149
DB 1086 GFOIGSEKMTGJIFISVAVSPAFPHSTI KPDKGLSVNSVSLDGVSHMALEIIQ 1145
QY 150 QAPKGLTREGALINKKIYISKT-SPV 177
DB 1146 NAEEDVT---LVISQPKRKISKVSTPV 1170

```

```

APPLICANT: FRANK, GLEN
TITLE OF INVENTION: VACCINE FOR SUSCEPTIBLE HOSTS AGAINST
TITLE OF INVENTION: NONADAPTED PARASITES
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESS: MORRISON & POERSTER
STREET: 755 PAGE MILL ROAD
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304-1018
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patulin Polcage #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/842,790
FILING DATE: 19920514
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: SHERA, DERRA A.
REGISTRATION NUMBER: 33,309
REFERENCE/WORKSHEET NUMBER: 27010-2000/00
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 813-5600
TELEFAX: (415) 494-0792
TELEX: 706141
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 428 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-882-790-4

Query Match 4.38; Score 79; DB 1; Length 428;
Best Local Similarity 21.14; Pred. No. 6;
Matches 75; Conservative 53; Mismatches 115; Indels 112; Gaps 20;

QY 71 FRKPRINVTIEFY--VYINDLFGNDN-----TSSVYVNGTPVWVY 121
DB 54 KKKAAE-----ELDGGCIMALSEIIQNEKMLMKIKISADPQIMKVEDMLKLVVD 107
QY 122 -----VPSFPHPTPPNST--HIVSSYSVH--WLDAPKGLTSP---ESLA 162
DB 108 KKKKFLIEFYVAFVFIYAMNFEFFENHLESYFTYLSWLTACFTEIKKKEEG-- 165
QY 163 LNKGIYISKTSPVVRNAVLSQHHDFITPLINARSOEYVPGCVLT--LRG----- 213
DB 166 --KSKMTGK-----ITVYPSLIGKKRAAEFGSPMAIPFVIGEPKVTI 214
QY 214 KQCSPEQDMQCEPMLLAMAIAELVSGLIDED--KLDP----- 252
DB 215 KQKMDP-----SAPRAELSMKVEHMF-KIVVQKPKKIDVAVPQKFAVIVHPPPHN 268
QY 253 --NIFSYF-----ASLENYGDIVEDGS-----FTIHIGSTLDP--SVEMQ 290
DB 269 FHNFSFYQTVYSWITPAKPEIKKKRPQKSKMTGKKRTPVVFSTIGPKKRAAEPIQ 328
QY 291 ENDKW-----VQGEKPTKVRVAETEPILSNQGPEDIMDKLYDKFTHIVSDJLAKL 341
DB 329 EGTMAIPFIVGFKWVILKQMDP-----SATIKELSMKVEHFKRIVVIRKPKKRI 360

```

```

RESULT 9
US-07-882-790 4
Sequence 4, Application US/07802750
Patent No. 5492695
GENERAL INFORMATION:
APPLICANT: ROBERT R

```

```

RESULT 10
US-08-540-804-14
Sequence 14, Application US/08540804
Patent No. 5919666
GENERAL INFORMATION:
APPLICANT: Koleske, Richard A.
APPLICANT: Koleske, Anthony J.

```

```

APPLICANT: Thompson, Craig M.
APPLICANT: Chao, David M.
TITLE OF INVENTION: No. 5919666e1 Factors Which Modify Gene
TITLE OF INVENTION: Transcription and Methods of Use Therefor
NUMBER OF SEQUENCES: 39
CORRESPONDENCE ADDRESS:
ADDRESSER: Hamilton, Brook, Smith & Reynolds, P.C.
STREET: Two Militia Drive
CITY: Lexington
STATE: Massachusetts
COUNTRY: USA
ZIP: 02173
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/540,804
FILING DATE: 11-OCT-1995
CLASSIFICATION: 424
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/521,872
FILING DATE: 21-AUG-1995
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/218,265
FILING DATE: 25-MAR-1994
ATTORNEY/AGENT INFORMATION:
NAME: Granahan, Patricia
REGISTRATION NUMBER: 32,227
REFERENCE/DOCKET NUMBER: WH194-03a2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-861-6240
TELEFAX: 617-861-9540
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 1420 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-540-804-14

```

Query Match 4.38; Score 79; DB 2; Length 1420;

Best Local Similarity 20.08; Pred. No. 41;

Matches 86; Conservative 56; Mismatches 161; Indels 126; Gaps 18;

```

10 SYAONSSTFOVASMOPALENAVETLFSRPHDIAIN----- 47
11 | | | | | | | | | | | | | | | | | | | | | |
33 SKAONDQMSIQMFELRKQDFTLVALLSRDLQCSINDDPYTPFALEHKVSPDKIGT 92
12 | | | | | | | | | | | | | | | | | | | | | |
48 AADLCAACPTFAVISTIKEMMEKKREINQTHLVVYINQFQNFENLEKGLSSE 106
13 | | | | | | | | | | | | | | | | | | | | | |
93 FTADYSKPNLPPIRYA-----LFLKALRKRIYNLAIGS-HNKLIO----- 131
14 | | | | | | | | | | | | | | | | | | | | | |
107 VTCNK-----FEVPGYVMV-PSSE-----HGLPPNSLIHVHSSYSVHMLTQ 150
15 | | | | | | | | | | | | | | | | | | | | | |
132 -FGNACTISGVYNTLVLEPHLPVNGDLTVSLCAKNMGLVPMKEENLEESLSKHALYL 190
16 | | | | | | | | | | | | | | | | | | | | | |
151 APKGLTSREGALNKGKRIYISKTPVREAVLSQFHEDFTMFL-----NARSQEVVNGC 206
17 | | | | | | | | | | | | | | | | | | | | | |
191 APGQIKHHLAPASKQGYL-----ITPKHTELLITLTVSHGINTQNKKNLKWAAVVPD-- 244
18 | | | | | | | | | | | | | | | | | | | | | |
207 MALLRPGQSPSPDMQSFTHWLLAMATAELVSGQLIDENKLDTPFNIPSPASLEEVKD 266
19 | | | | | | | | | | | | | | | | | | | | | |
245 -ESHLNQHFTTASVLTPLLEAKKLVWPLHLIFAQFVADIEN-STSGDSEFHCLQDALD 302
20 | | | | | | | | | | | | | | | | | | | | | |
267 IVE-----RDGST--IDHLEPGDDLSV-----E 288
21 | | | | | | | | | | | | | | | | | | | | | |
303 AIDDFIOLKOTAAVVRPGSSCVLSNIACTNPLSSDQATTEOFQHYKKNNSISSQASYHS 362
22 | | | | | | | | | | | | | | | | | | | | | |
289 MOENDWVRGE---KFTKVRAFTPELISNOCPELM-----DKLY-DKETHIAYS 335
23 | | | | | | | | | | | | | | | | | | | | | |

```

Db 363 VQENKISPKTFSPNFTGRTIKIMSP--STGFAPAFNTPNNNINPNEIENIKQKQTVSN 420

QY 336 DLEAKLPKT 344

Db 421 DLENSLPKT 429

RESULT 11

US-08-218-265-14

Sequence 14, Application US/08218265

Patent No. 5922585

GENERAL INFORMATION:

APPLICANT: Young, Richard A.

APPLICANT: Koleske, Anthony J.

TITLE OF INVENTION: No. 5922585e1 Factors Which Modify Gene

TITLE OF INVENTION: Transcription and Methods of Use Thereof

NUMBER OF SEQUENCES: 35

CORRESPONDENCE ADDRESS:

ADDRESSER: Hamilton, Brook, Smith & Reynolds, P.C.

STREET: Two Militia Drive

CITY: Lexington

STATE: MA

COUNTRY: US

ZIP: 02173

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/218,265

FILING DATE: 25-MAR-1994

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Granahan, Patricia

REGISTRATION NUMBER: 32,227

REFERENCE/DOCKET NUMBER: WH194-03

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617-861-6240

TELEFAX: 617-861-9540

INFORMATION FOR SEQ ID NO: 14:

SEQUENCE CHARACTERISTICS:

LENGTH: 1420 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-218-265-14

Query Match 4.38; Score 79; DB 2; Length 1420;

Best Local Similarity 20.08; Pred. No. 41;

Matches 86; Conservative 56; Mismatches 161; Indels 126; Gaps 18;

```

10 SYAONSSTFOVASMOPALENAVETLFSRPHDIAIN----- 47
11 | | | | | | | | | | | | | | | | | | | | | |
33 SKAONDQMSIQMFELRKQDFTLVALLSRDLQCSINDDPYTPFALEHKVSPDKIGT 92
12 | | | | | | | | | | | | | | | | | | | | | |
48 AADLCAACPTFAVISTIKEMMEKKREINQTHLVVYINQFQNFENLEKGLSSE 106
13 | | | | | | | | | | | | | | | | | | | | | |
93 FTADYSKPNLPPIRYA-----LFLKALRKRIYNLAIGS-HNKLIO----- 131
14 | | | | | | | | | | | | | | | | | | | | | |
107 VTCNK-----FEVPGYVMV-PSSE-----HGLPPNSLIHVHSSYSVHMLTQ 150
15 | | | | | | | | | | | | | | | | | | | | | |
132 -FGNACTISGVYNTLVLEPHLPVNGDLTVSLCAKNMGLVPMKEENLEESLSKHALYL 190
16 | | | | | | | | | | | | | | | | | | | | | |
151 APKGLTSREGALNKGKRIYISKTPVREAVLSQFHEDFTMFL-----NARSQEVVNGC 206
17 | | | | | | | | | | | | | | | | | | | | | |
191 APGQIKHHLAPASKQGYL-----ITPKHTELLITLTVSHGINTQNKKNLKWAAVVPD-- 244
18 | | | | | | | | | | | | | | | | | | | | | |
207 MALLRPGQSPSPDMQSFTHWLLAMATAELVSGQLIDENKLDTPFNIPSPASLEEVKD 266
19 | | | | | | | | | | | | | | | | | | | | | |
245 -ESHLNQHFTTASVLTPLLEAKKLVWPLHLIFAQFVADIEN-STSGDSEFHCLQDALD 302
20 | | | | | | | | | | | | | | | | | | | | | |
267 IVE-----RDGST--IDHLEPGDDLSV-----E 288
21 | | | | | | | | | | | | | | | | | | | | | |
303 AIDDFIOLKOTAAVVRPGSSCVLSNIACTNPLSSDQATTEOFQHYKKNNSISSQASYHS 362
22 | | | | | | | | | | | | | | | | | | | | | |
245 -ESHLNQHFTTASVLTPLLEAKKLVWPLHLIFAQFVADIEN-STSGDSEFHCLQDALD 302
23 | | | | | | | | | | | | | | | | | | | | | |

```

QY 267 IV-----RGSFT--IDHTEGDLDSV-----E 288  
 DB 303 AIDFIOIKOTAAVYKPGSSGLVSSNAGTNPPLSSDGAATFQYHKNNSSOPASYHS 362  
 QY 289 MGENKWKWRC---KTKVFAKTEPTISNQGEPFIM-----DKLY-DKTHIVYS 335  
 DB 363 VQFNKISPKDFSNFTGTRKIMLSF--SDGFAPALNTPNNINENELFNPKQTTYSN 420  
 QY 336 DLEAKLPKT 344  
 DB 421 DLENSPLKT 429

RESULT 12  
 US-08-521-872-14  
 : Sequence 14, Application US/08/521872  
 : Patent No. 6015682  
 : GENERAL INFORMATION:  
 : APPLICANT: Young, Richard A.  
 : APPLICANT: Koleske, Anthony J.  
 : APPLICANT: Thompson, Craig M.  
 : APPLICANT: Chao, David M.  
 : TITLE OF INVENTION: No. 6015682 Factors Which Modify Gene  
 : TITLE OF INVENTION: Transcription and Methods of Use Therefor  
 : NUMBER OF SEQUENCES: 37  
 : CORRESPONDENCE ADDRESS:  
 : ADDRESSEE: Hamilton, Henry, Smith & Reynolds, P.C.  
 : STREET: Two Militia Drive  
 : CITY: Lexington  
 : STATE: Massachusetts  
 : COUNTRY: USA  
 : ZIP: 02173  
 : COMPUTER READABLE FORM:  
 : MEDIUM TYPE: Floppy disk  
 : COMPUTER: IBM PC compatible  
 : OPERATING SYSTEM: PC-DOS/MS-DOS  
 : SOFTWARE: Patent in Release #1.0, Version #1.30  
 : CURRENT APPLICATION DATA:  
 : APPLICATION NUMBER: US/08/521,872  
 : FILING DATE: 31-AUG-1995  
 : CLASSIFICATION: 436  
 : PRIOR APPLICATION DATA:  
 : APPLICATION NUMBER: US 08/218,265  
 : FILING DATE: 25-MAR-1994  
 : AUTOPNRY/AGENT INFORMATION:  
 : NAME: Granahan, Patricia  
 : REGISTRATION NUMBER: 32,227  
 : REFERENCE/AGENT NUMBER: WH194-03A3  
 : TELECOMMUNICATION INFORMATION:  
 : TELEPHONE: 617-861-9540  
 : TELEFAX: 617-861-9540  
 : INFORMATION FOR SEQ ID NO: 14:  
 : SEQUENCE CHARACTERISTICS:  
 : LENGTH: 1420 amino acids  
 : TYPE: amino acid  
 : TOPOLOGY: linear  
 : MOLECULE TYPE: protein  
 : US-08-521-872-14

Query Match 4 8, Score 79, P93, Length 1420,  
 Best Local Similarity 20 08; P93, No. 41;  
 Matches 86; Conservative 56; Mismatches 161; Indels 126; Gaps 18;  
 QY 10 SYAONSSTQGVASMAQPALENAFTLSKPFHLDAN----- 47  
 DB 33 SKAONDWSIOMELMKKQKTLVALLSKRWESINIDWVPIPAALPKKVSPIKIC 92  
 QY 48 -AALDGAGAGNFAVATISIKRMMEKKQELNVTLELVYLMLEPNFNILFKGISSE 106  
 DB 93 FTAYSKRNLPPIHYA-----LPLKALPRKITYINLAIGS-HNKLIG----- 131

QY 107 VIGNKC---EEVRYVWGV-PGSP-----HGRIIPRNSIHLVHSSVYHMTQ 150  
 DB 132 -FQNACTISGVPNVYVQLEPFLPVN-DLTVSLQAKNMGIYMKKEFNLEFSLSKHALYL 190  
 QY 151 AKGLTSPECIALNKCKIKYISKTSPPVREAVYSQFHEFTMTI---NAPSGEVVWPG 206  
 DB 191 APSGIRHILAPASKGTL---TTPKHTELLITL SVSHGTLNZNKKNLKVAAVVP-- 244  
 QY 207 MYILPGPGSDPSQSGFTWEILAMAIAYVSGQIDENKLTFFNIPSFASLEEKD 266  
 DB 245 -LGLNCHFTPLIASYLPLEAKKLWPLHLIFAQVADIEN-STSGDSEPHQLQDALD 302  
 QY 267 IV-----RGSFT--IDHTEGDLDSV-----E 288  
 DB 303 AIDFIOIKOTAAVYKPGSSGLVSSNAGTNPPLSSDGAATFQYHKNNSSOPASYHS 362  
 QY 289 MGENKWKWRC---KTKVFAKTEPTISNQGEPFIM-----DKLY-DKTHIVYS 335  
 DB 363 VQFNKISPKDFSNFTGTRKIMLSF--SDGFAPALNTPNNINENELFNPKQTTYSN 420  
 QY 336 DLEAKLPKT 344  
 DB 421 DLENSPLKT 429

RESULT 13  
 US-08-590-399-14  
 : Sequence 14, Application US/08/590399  
 : Patent No. 6214588  
 : GENERAL INFORMATION:  
 : APPLICANT: Young, Richard A.  
 : APPLICANT: Koleske, Anthony J.  
 : APPLICANT: Thompson, Craig M.  
 : APPLICANT: Chao, David M.  
 : TITLE OF INVENTION: No. 6214588 Factors Which Modify Gene  
 : TITLE OF INVENTION: Transcription and Methods of Use Therefor  
 : NUMBER OF SEQUENCES: 39  
 : CORRESPONDENCE ADDRESS:  
 : ADDRESSEE: Hamilton, Henry, Smith & Reynolds, P.C.  
 : STREET: Two Militia Drive  
 : CITY: Lexington  
 : STATE: Massachusetts  
 : COUNTRY: USA  
 : ZIP: 02173  
 : COMPUTER READABLE FORM:  
 : MEDIUM TYPE: Floppy disk  
 : COMPUTER: IBM PC compatible  
 : OPERATING SYSTEM: PC-DOS/MS-DOS  
 : SOFTWARE: Patent in Release #1.0, Version #1.30  
 : CURRENT APPLICATION DATA:  
 : APPLICATION NUMBER: US/08/590,399  
 : FILING DATE: 26-JAN-1996  
 : CLASSIFICATION: 435  
 : PRIOR APPLICATION DATA:  
 : APPLICATION NUMBER: US 08/540,804  
 : FILING DATE: 11-OCT-1995  
 : PRIOR APPLICATION DATA:  
 : APPLICATION NUMBER: US 08/521,872  
 : FILING DATE: 31-AUG-1995  
 : AUTOPNRY/AGENT INFORMATION:  
 : NAME: Granahan, Patricia  
 : REGISTRATION NUMBER: 32,227  
 : REFERENCE/AGENT NUMBER: WH194-03A3  
 : TELECOMMUNICATION INFORMATION:  
 : TELEPHONE: 617-861-9540  
 : TELEFAX: 617-861-9540  
 : INFORMATION FOR SEQ ID NO: 14:  
 : SEQUENCE CHARACTERISTICS:  
 : LENGTH: 1420 amino acids  
 : TYPE: amino acid

TOPOLGY: linear  
MOLECULE TYPE: protein  
US 08-590-399-14

Query Match: 4.78, Score 79, DB 4, Length 1420,  
Best Local Similarity 20.0%, Pred. No. 41;  
Matches 86; Conservative 56; Mismatches 161; Indels 126; Gaps 18.

```

UY 10 SYAONSSFTQVAVSMAOPALENAVELFESRDLQALN----- 47
DB 33 SKAONPMWSIQMEPMKPKPLVALISPIWESINDEPVPFVALHKHVSPOKIGT 92
UY 48 -AALDCAAPNTEFAVISTIKRMKEKKCELOCOLEQVINDLEFGDNFTLFGLSSE 106
DB 93 FTADYSKPNLPPIYA-----LPIKALRKITYINLAIGS HNKLTG----- 131
UY 107 VIGNKC---EVNCCYMGV PGSE-----HGRLEPRNSLHLVSSVYHWLTQ 150
DB 132 -FPMACISLGVNRYVQLEPHLFVWQDLTVSLCAKNMGLVPMKEFNEESLSKHALYL 190
UY 151 APKITTEBETATNKKITYISKISFVVPFATISCHETFTMFL-----NAPSELEVVPKQ 206
DB 191 APSGIRMHAPASKQGYL-----ITPKHTELTLTILSVSHGINLKNKKIKWAAVVPD-- 244
UY 207 MVLLRQKQSTPESFMSCFTWELLAMALAEVSGLIHEKLEFNLSPYASLEFVKD 266
DB 245 LCHLNKHTITLASYTLTLLAKKILVWPLHLIFAGVVALEN-SISQISEPHLQALD 302
UY 267 IVE-----PSTFT--IHLLEGLLSV-----E 288
DB 303 AIDDFIOLKOTAAVKTGSSNVLSSNIAGTNPSSDQAVTEQFQHKNSISSOPASYHS 362
UY 289 MOENDKAVRGE---KTKVYKATFTEPISNPGPELM-----DKLY-DKPFHIVS 335
DB 363 VQENKISPEKDFSPNTGIDKMLSP--SDQAPAFNLPPNNINDELNDKQOTVSN 420
UY 336 DEAKLPKT 344
DB 421 DEENSPKLT 429

RESULT 14
US-08-286-819A-27
Sequence 27, Application US-08-286-819A
Patent No. 5871910
GENERAL INFORMATION:
APPLICANT: ARTHUR, MICHEL
APPLICANT: PIKTA, MALEN, SYLVIE
APPLICANT: MOLINAS, CATHERINE
TITLE OF INVENTION: POLYPEPTIDES IMPLICATED IN THE
TITLE OF INVENTION: EXPRESSION OF PESTICANT TO GLYCOPETIDES, IN PARTICULAR
TITLE OF INVENTION: IN GRAM-POSITIVE BACTERIA, NUCLEOTIDE SEQUENCE CODING FOR
NUMBER OF SEQUENCES: 54
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
ADDRESSEE: P.C.
STREET: 1765 S. Jefferson Davis Highway, Suite 400
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US-08/286,819A
FILING DATE: 05-AUG-1994
CLASSIFICATION: 435

```

```

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/174,682
FILING DATE: 28-DEC-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/917,146
FILING DATE: 10-AUG-1992
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/FR/91/200855
FILING DATE: 29-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 9013579
FILING DATE: 31-OCT-1990
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Oblon, No. 5871910man F.
REGISTRATION NUMBER: 24,618
REFERENCE/WORKER NUMBER: no-000-0 PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 413-4000
TELEFAX: (703) 413-2250
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 27:
SEQUENCE CHARACTERISTICS:
LENGTH: 2296 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-286-819A-27

Query Match: 4.3%, Score 78.5, DB 2, Length 2296,
Best Local Similarity 18.0%, Pred. No. 99;
Matches 70; Conservative 68; Mismatches 103; Indels 147; Gaps 22;

UY 53 CAAGPNTFAVIST-----IKRME-----KKC---RLN---CUTDEL----- 84
DB 241 GQDFPESFKTISAVDRDNGYIKRPFQCGKIASNPKCTIYPMKTSFMLQTIESLVH 300
UY 85 QVTLNDLPGNDFNTLFGKLSSEV---IGKCKEVCYMGVGSFGKLEPRNSLHLVHS 141
DB 301 KYLEN-----SLYEIKRQVNR-----IHGLYGR-----IHT 340
UY 142 SYVHMLIAKELISEH-----ALR---FKKITYISKISFVVEEAVLSQEHEDFT 191
DB 331 ANSLMLV-----LTNSLIFPNVWQINMLKKKESIKVPSIITGTFIMPKELYEKGN 385
UY 192 MLN-----APSEVFNQCMVLLKQF---QSTQVPMSCFTW 228
DB 386 MYNQDEITNVSFASLYKLSSEV--NNPFCPLDMJNLFKEKELIKYQPSMAEPPDK 444
UY 429 ELIAMA-----LALVSGQILDEDKIDTENTISYFAS-----EYKQIVE R 270
DB 444 KPTILLNSSEFIMMLNSA--SPKLLPANICEVYTSQKLLVDEBHHADLVEYLK 501
UY 271 DGSEFT-----DHIEGPNDSVFMQENQVWPCEKTYVVPATETISNPGPIKQ 324
DB 502 NENYIVFKYITAKFALECIDKSEIDL-----AIDIMIGTSGITG 544
UY 324 KLYDKTH--IYVSDLEAKLPKTTSTIL 349
DB 545 KIRDKHTYPLMKQDKVWKITGLTI 572

RESULT 15
US-08-980-357-27
Sequence 27, Application US-08-980-357
Patent No. 6013508
GENERAL INFORMATION:
APPLICANT: ARTHUR, MICHEL
APPLICANT: PIKTA, MALEN, SYLVIE
APPLICANT: MOLINAS, CATHERINE

```

```

APPLICANT: COBRALIN, PATRICE
TITLE OF INVENTION: POLYPEPTIDES IMPLICATED IN THE
TITLE OF INVENTION: EXPRESSION OF RESISTANCE TO GLYCOPOLYMER
TITLE OF INVENTION: IN GRAM POSITIVE BACTERIA, NUCLEOTIDE SEQUENCE
TITLE OF INVENTION: THESE POLYPEPTIDES AND USE FOR DIAGNOSIS
NUMBER OF SEQUENCES: 54
CORRESPONDENCE ADDRESS:
ADDRESS: P.O.
ADDRESS: GILSON, SPILVER, MONTILLANT, MAIRIE & NEUSILLY
STREET: 1755 S. Jefferson Davis Highway, Suite 400
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/980,457
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/286,819
FILING DATE: 05-AUG-1994
APPLICATION NUMBER: US 08/174,682
FILING DATE: 28-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/917,146
FILING DATE: 10-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/FR/91/00855
FILING DATE: 29-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 9013579
FILING DATE: 31-OCT-1990
ATTORNEY/AGENT INFORMATION:
NAME: Obion, No. 6013508man P.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 660-060-0 PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 413-3000
TELEFAX: (703) 413-2220
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 27:
SEQUENCE CHARACTERISTICS:
LENGTH: 2296 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-980-357-27

```

```

Query Match 4 % Score 78 5; PR 3; length 2296;
Best local similarity 18.0% Pred. No. 99;
Matches 70; Conservative 68; Mismatches 103; Indels 147; Gaps 22;

```

```

QY 53 CAAGPNTFAVIST-----IKRME-----KRC---RELN---COTLEL----- 84
Db 241 CQTPPEFSAKTSIAVEKRNQYVIAKRFQEDQRIASNEKKQIYKMMISMLQTIISLVVH 300
QY 85 QVYINDLEFGNDPNTLEFGKLSSEV---IGNKCEVPCYVMGVPGSGFGRLEPRNSLTVHS 141
Db 301 KIYLN-----SITYEIKROYNHK-----IHGLIYQK-----IHT 330
QY 142 SYSVHMLTQAPKGLTSREGL-----ALN---GKTIYISKISPPVVRKAYLSQFHEOPT 191
Db 331 ANSLMLV-----LTNSEILFGDNVKGMLNLRKRESIKVDPSITIKITQELMRKLYKEGN 385
QY 192 MFLN-----ARSGFVPPNCGVALLPGR-----QSNPSMOSCFPW 228
Db 386 MIVMOIGETITNVSRASLYRKLSEV--NNPFCIPLMGNIFKEKEKITKYQPPSDAEKPFDK 443

```

```

QY 229 EILAMA - IAEIVSQGLLEDAIDITNIESYVSI-----EYKDIY---P 270
Db 444 KELLINSSSEIMMLINSAL--SKFELPANIYEVYIMSEKIIIMVGHFIADLEVETLK 501
QY 271 WSEFTI-----FHRCPTLDSVEPMQPNKRWPGCRKPTKRVKATFTTITSNGPSPFIMD 323
Db 502 NENIYVHYIAEAGLCILKSEID-----AIDIMLPISGLTIQ 544
QY 424 KLYDKFTH--IWSDBAKLPKITSIL 349
Db 545 KIRDKHTYPIIMLGKDTEDKITGLTI 572

```

```

Search completed: August 31, 2001, 15:00:28
Job time: 2161 sec

```

.  
.  
.

Genome version 4.5  
Copyright (c) 1993 - 2000 Computer Ltd.

OM protein - protein search, using SW model

Run on: August 31, 2001 14:22:22 Search time 45.19 seconds  
(without alignments)  
477.587 Million cell updates/sec

Title: US-09-577-657A-1

Perfect score: 1847

Sequence: 1 FMNPFGRSSVAQNSSTQQ

Scoring table: BLOSUM62

Gapop 10.0 Gapext 0.5

Searched: 412674 seqs, 7862998 residues

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Maximum Match 09

Listing first 45 summaries

Database: A\_Geneseq\_0901.\*

1: /SIDSI/gcgdata/geneseq/geneseq/AA1980.DAT.\*

2: /SIDSI/gcgdata/geneseq/geneseq/AA1981.DAT.\*

3: /SIDSI/gcgdata/geneseq/geneseq/AA1982.DAT.\*

4: /SIDSI/gcgdata/geneseq/geneseq/AA1983.DAT.\*

5: /SIDSI/gcgdata/geneseq/geneseq/AA1984.DAT.\*

6: /SIDSI/gcgdata/geneseq/geneseq/AA1985.DAT.\*

7: /SIDSI/gcgdata/geneseq/geneseq/AA1986.DAT.\*

8: /SIDSI/gcgdata/geneseq/geneseq/AA1987.DAT.\*

9: /SIDSI/gcgdata/geneseq/geneseq/AA1988.DAT.\*

10: /SIDSI/gcgdata/geneseq/geneseq/AA1989.DAT.\*

11: /SIDSI/gcgdata/geneseq/geneseq/AA1990.DAT.\*

12: /SIDSI/gcgdata/geneseq/geneseq/AA1991.DAT.\*

13: /SIDSI/gcgdata/geneseq/geneseq/AA1992.DAT.\*

14: /SIDSI/gcgdata/geneseq/geneseq/AA1993.DAT.\*

15: /SIDSI/gcgdata/geneseq/geneseq/AA1994.DAT.\*

16: /SIDSI/gcgdata/geneseq/geneseq/AA1995.DAT.\*

17: /SIDSI/gcgdata/geneseq/geneseq/AA1996.DAT.\*

18: /SIDSI/gcgdata/geneseq/geneseq/AA1997.DAT.\*

19: /SIDSI/gcgdata/geneseq/geneseq/AA1998.DAT.\*

20: /SIDSI/gcgdata/geneseq/geneseq/AA1999.DAT.\*

21: /SIDSI/gcgdata/geneseq/geneseq/AA2000.DAT.\*

22: /SIDSI/gcgdata/geneseq/geneseq/AA2001.DAT.\*

Prod. No is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1847	100.0	356	22	AAH45799
2	716.5	38.8	386	21	AAH45799
3	716.5	38.8	386	21	AAH45799
4	716.5	38.8	405	21	AAH45799
5	687	37.2	349	21	AAH45799
6	687	37.2	349	21	AAH45799
7	687	37.2	379	21	AAH45799
8	687	37.2	379	21	AAH45799
9	641	34.7	327	21	AAH45799
10	641	34.7	327	21	AAH45799
11	481	26.0	374	21	AAH45799

12	481	26.0	386	21	AAH45799	Arabidopsis thaliana
13	481	26.0	387	21	AAH45799	Arabidopsis thaliana
14	477	25.8	374	21	AAH45799	Arabidopsis thaliana
15	477	25.8	386	21	AAH45799	Arabidopsis thaliana
16	477	25.8	386	21	AAH45799	Arabidopsis thaliana
17	446.5	24.2	348	21	AAH45799	Arabidopsis thaliana
18	446	24.1	343	21	AAH45799	Arabidopsis thaliana
19	415	22.5	345	21	AAH45799	Arabidopsis thaliana
20	415	22.5	345	21	AAH45799	Arabidopsis thaliana
21	364	18.8	318	21	AAH45799	Arabidopsis thaliana
22	347	18.8	255	21	AAH45799	Arabidopsis thaliana
23	319	17.3	181	21	AAH45799	Arabidopsis thaliana
24	319	17.3	187	21	AAH45799	Arabidopsis thaliana
25	145.5	7.9	154	21	AAH45799	Arabidopsis thaliana
26	140.5	7.6	117	21	AAH45799	Arabidopsis thaliana
27	140	7.6	112	21	AAH45799	Arabidopsis thaliana
28	97.5	4.3	274	18	AAH45799	Arabidopsis thaliana
29	92.5	5.0	751	14	AAH45799	Arabidopsis thaliana
30	92.5	5.0	751	14	AAH45799	Arabidopsis thaliana
31	92.5	5.0	335	17	AAH45799	Arabidopsis thaliana
32	91.5	4.9	412	12	AAH45799	Arabidopsis thaliana
33	91	4.9	446	21	AAH45799	Arabidopsis thaliana
34	91	4.9	446	21	AAH45799	Arabidopsis thaliana
35	90	4.9	446	18	AAH45799	Arabidopsis thaliana
36	89.5	4.8	568	21	AAH45799	Arabidopsis thaliana
37	87	4.7	20	22	AAH45799	Arabidopsis thaliana
38	87	4.7	26	21	AAH45799	Arabidopsis thaliana
39	87	4.7	823	21	AAH45799	Arabidopsis thaliana
40	87	4.7	954	21	AAH45799	Arabidopsis thaliana
41	85.5	4.7	1447	22	AAH45799	Arabidopsis thaliana
42	84	4.7	447	21	AAH45799	Arabidopsis thaliana
43	84	4.7	447	21	AAH45799	Arabidopsis thaliana
44	84	4.7	440	21	AAH45799	Arabidopsis thaliana
45	84	4.7	444	21	AAH45799	Arabidopsis thaliana

#### ALIGNMENTS

RESULT 1

AAH45799 standard: protein: 356 AA.

XX

XX AAH45799:

XX

XX 16-MAR-2001 (first entry)

XX

XX C. sinensis N-methyl transferase protein.

XX

XX Caffeine biosynthesis, N-methyl transferase, tea, N-methyl plant.

XX

XX Camellia sinensis.

XX

XX EP1055727-A2.

XX

XX 29-NOV-2000.

XX

XX 26-MAR-2000; 2000EP-0304522

XX

XX 26-MAR-1999; 99UP-0146358.

XX

XX (MITA) MITSUI CHEM INC.

XX

XX Mizuno M, Ashihara H, Mizuno K, Fujimura T;

XX

XX Wt1; 2001-06-01/08

XX

XX RNA encoding N-methyl transferase with enzyme activities of 7-methyl

XX

XX xanthine N3-methyl transferase, theobromine N1-methyl transferase and

XX

XX paraxanthine N3-methyl transferase, useful for caffeine synthesis in

XX

XX plants or microorganisms -

XX

XX Claim 1; page 19; 33pp; English.

```

XX  This invention provides a novel DNA molecule (1) encoding N-methyl
CC transferase (N-met) with the enzyme activities of 7-methyl xanthine N3
CC methyl transferase, theobromine N-methyl transferase and paraxanthine
CC N3 methyl transferase. The DNA or RNA encoding the N-met is useful for
CC the enhancement of caffeine production in microorganisms or plants, and
CC the encoded N-met is also used for caffeine synthesis. The methods and
CC the DNA may also be used for producing N-met that can be used as an
CC enzyme for industrial, food or medical use, e.g. for producing compounds
CC related to caffeine metabolism by modifying caffeine biosynthesis and
CC metabolism of caffeine productive plants, plant tissues or cells and for
CC modifying caffeine biosynthesis and metabolism of caffeine productive
CC plants, plant tissues or cells, therefore modifying the production rate
CC of caffeine metabolism related compounds.
XX
SU Sequence 356 AA:

Query Match 100.0%; Score 1847; E-Value 2.2; Length 356;
Best Local Similarity 100.0%; Ident. No. 2,76-177;
Matches 356; Conservative 0; Mismatches 0; Indels 0; Gaps 0.

UY 1 FMNKGREESYAGNSSTQGVASMAQALBNAYETLPSKDHUQALNADLCAAGPNTF 60
    |||||||
Db 1 FMNKGREESYAGNSSTQGVASMAQALBNAYETLPSKDHUQALNADLCAAGPNTF 60

UY 61 AVISTEKRMKEKKKPEI N7GTEFLGVYNIENFENENILEKGLSSSEVLNAGEEPCVYM 120
    |||||||
Db 61 AVISTEKRMKEKKKPEI N7GTEFLGVYNIENFENENILEKGLSSSEVLNAGEEPCVYM 120

UY 121 GVPQSEHGLPEPNSI HLHSSYSVHMLTQAKGTTSPGLANRQKYLSTSPVVPF 180
    |||||||
Db 121 GVPQSEHGLPEPNSI HLHSSYSVHMLTQAKGTTSPGLANRQKYLSTSPVVPF 180

UY 181 AYL SQEPFETMEL NARSGVAVHNCVMVLTLPQPCSPNSQSGPTFWILLAMAFAELVS 240
    |||||||
Db 181 AYL SQEPFETMEL NARSGVAVHNCVMVLTLPQPCSPNSQSGPTFWILLAMAFAELVS 240

UY 241 QGLDDEKDKLTFTNPSYASLEEKDIVERKSFETTHIESFELASVEKEMKNKVPGRK 300
    |||||||
Db 241 QGLDDEKDKLTFTNPSYASLEEKDIVERKSFETTHIESFELASVEKEMKNKVPGRK 300

UY 301 FTWVAVAFPTPLISNQGVEHIMKLYDKFTHIVSLELEKPKTISLIVLSKIDG 356
    |||||||
Db 301 FTWVAVAFPTPLISNQGVEHIMKLYDKFTHIVSLELEKPKTISLIVLSKIDG 356

RESULT 2
AA122190
ID AAG22190 standard; protein; 386 AA.
XX
AC AAG22190:
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 25024
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1013405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000: 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825
PR 05-MAR-1999; 99US-0128180.
PR 09-MAR-1999; 99US-0123548.
PR 21-MAR-1999; 99US-0125788.

```

```

PR 25-MAR-1999; 99US-0126364.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128334.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129445.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130810.
PR 25-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 18-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138040.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139464.
PR 18-JUN-1999; 99US-0139465.
PR 18-JUN-1999; 99US-0139466.
PR 18-JUN-1999; 99US-0139467.
PR 18-JUN-1999; 99US-0139468.
PR 18-JUN-1999; 99US-0139469.
PR 18-JUN-1999; 99US-0139470.
PR 18-JUN-1999; 99US-0139471.
PR 18-JUN-1999; 99US-0139472.
PR 18-JUN-1999; 99US-0139473.
PR 18-JUN-1999; 99US-0139474.
PR 18-JUN-1999; 99US-0139475.
PR 18-JUN-1999; 99US-0139476.
PR 18-JUN-1999; 99US-0139477.
PR 18-JUN-1999; 99US-0139478.
PR 18-JUN-1999; 99US-0139479.
PR 18-JUN-1999; 99US-0139480.
PR 18-JUN-1999; 99US-0139481.
PR 18-JUN-1999; 99US-0139482.
PR 18-JUN-1999; 99US-0139483.
PR 18-JUN-1999; 99US-0139484.
PR 18-JUN-1999; 99US-0139485.
PR 18-JUN-1999; 99US-0139486.
PR 18-JUN-1999; 99US-0139487.
PR 18-JUN-1999; 99US-0139488.
PR 18-JUN-1999; 99US-0139489.
PR 18-JUN-1999; 99US-0139490.
PR 18-JUN-1999; 99US-0139491.
PR 18-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139493.
PR 18-JUN-1999; 99US-0139494.
PR 18-JUN-1999; 99US-0139495.
PR 18-JUN-1999; 99US-0139496.
PR 18-JUN-1999; 99US-0139497.
PR 18-JUN-1999; 99US-0139498.
PR 18-JUN-1999; 99US-0139499.
PR 18-JUN-1999; 99US-0139500.
PR 18-JUN-1999; 99US-0139501.
PR 18-JUN-1999; 99US-0139502.
PR 18-JUN-1999; 99US-0139503.
PR 18-JUN-1999; 99US-0139504.
PR 18-JUN-1999; 99US-0139505.
PR 18-JUN-1999; 99US-0139506.
PR 18-JUN-1999; 99US-0139507.
PR 18-JUN-1999; 99US-0139508.
PR 18-JUN-1999; 99US-0139509.
PR 18-JUN-1999; 99US-0139510.
PR 18-JUN-1999; 99US-0139511.
PR 18-JUN-1999; 99US-0139512.
PR 18-JUN-1999; 99US-0139513.
PR 18-JUN-1999; 99US-0139514.
PR 18-JUN-1999; 99US-0139515.
PR 18-JUN-1999; 99US-0139516.
PR 18-JUN-1999; 99US-0139517.
PR 18-JUN-1999; 99US-0139518.
PR 18-JUN-1999; 99US-0139519.
PR 18-JUN-1999; 99US-0139520.
PR 18-JUN-1999; 99US-0139521.
PR 18-JUN-1999; 99US-0139522.
PR 18-JUN-1999; 99US-0139523.
PR 18-JUN-1999; 99US-0139524.
PR 18-JUN-1999; 99US-0139525.
PR 18-JUN-1999; 99US-0139526.
PR 18-JUN-1999; 99US-0139527.
PR 18-JUN-1999; 99US-0139528.
PR 18-JUN-1999; 99US-0139529.
PR 18-JUN-1999; 99US-0139530.
PR 18-JUN-1999; 99US-0139531.
PR 18-JUN-1999; 99US-0139532.
PR 18-JUN-1999; 99US-0139533.
PR 18-JUN-1999; 99US-0139534.
PR 18-JUN-1999; 99US-0139535.
PR 18-JUN-1999; 99US-0139536.
PR 18-JUN-1999; 99US-0139537.
PR 18-JUN-1999; 99US-0139538.
PR 18-JUN-1999; 99US-0139539.
PR 18-JUN-1999; 99US-0139540.
PR 18-JUN-1999; 99US-0139541.
PR 18-JUN-1999; 99US-0139542.
PR 18-JUN-1999; 99US-0139543.
PR 18-JUN-1999; 99US-0139544.
PR 18-JUN-1999; 99US-0139545.
PR 18-JUN-1999; 99US-0139546.
PR 18-JUN-1999; 99US-0139547.
PR 18-JUN-1999; 99US-0139548.
PR 18-JUN-1999; 99US-0139549.
PR 18-JUN-1999; 99US-0139550.
PR 18-JUN-1999; 99US-0139551.
PR 18-JUN-1999; 99US-0139552.
PR 18-JUN-1999; 99US-0139553.
PR 18-JUN-1999; 99US-0139554.
PR 18-JUN-1999; 99US-0139555.
PR 18-JUN-1999; 99US-0139556.
PR 18-JUN-1999; 99US-0139557.
PR 18-JUN-1999; 99US-0139558.
PR 18-JUN-1999; 99US-0139559.
PR 18-JUN-1999; 99US-0139560.
PR 18-JUN-1999; 99US-0139561.
PR 18-JUN-1999; 99US-0139562.
PR 18-JUN-1999; 99US-0139563.
PR 18-JUN-1999; 99US-0139564.
PR 18-JUN-1999; 99US-0139565.
PR 18-JUN-1999; 99US-0139566.
PR 18-JUN-1999; 99US-0139567.
PR 18-JUN-1999; 99US-0139568.
PR 18-JUN-1999; 99US-0139569.
PR 18-JUN-1999; 99US-0139570.
PR 18-JUN-1999; 99US-0139571.
PR 18-JUN-1999; 99US-0139572.
PR 18-JUN-1999; 99US-0139573.
PR 18-JUN-1999; 99US-0139574.
PR 18-JUN-1999; 99US-0139575.
PR 18-JUN-1999; 99US-0139576.
PR 18-JUN-1999; 99US-0139577.
PR 18-JUN-1999; 99US-0139578.
PR 18-JUN-1999; 99US-0139579.
PR 18-JUN-1999; 99US-0139580.
PR 18-JUN-1999; 99US-0139581.
PR 18-JUN-1999; 99US-0139582.
PR 18-JUN-1999; 99US-0139583.
PR 18-JUN-1999; 99US-0139584.
PR 18-JUN-1999; 99US-0139585.
PR 18-JUN-1999; 99US-0139586.
PR 18-JUN-1999; 99US-0139587.
PR 18-JUN-1999; 99US-0139588.
PR 18-JUN-1999; 99US-0139589.
PR 18-JUN-1999; 99US-0139590.
PR 18-JUN-1999; 99US-0139591.
PR 18-JUN-1999; 99US-0139592.
PR 18-JUN-1999; 99US-0139593.
PR 18-JUN-1999; 99US-0139594.
PR 18-JUN-1999; 99US-0139595.
PR 18-JUN-1999; 99US-0139596.
PR 18-JUN-1999; 99US-0139597.
PR 18-JUN-1999; 99US-0139598.
PR 18-JUN-1999; 99US-0139599.
PR 18-JUN-1999; 99US-0139600.
PR 18-JUN-1999; 99US-0139601.
PR 18-JUN-1999; 99US-0139602.
PR 18-JUN-1999; 99US-0139603.
PR 18-JUN-1999; 99US-0139604.
PR 18-JUN-1999; 99US-0139605.
PR 18-JUN-1999; 99US-0139606.
PR 18-JUN-1999; 99US-0139607.
PR 18-JUN-1999; 99US-0139608.
PR 18-JUN-1999; 99US-0139609.
PR 18-JUN-1999; 99US-0139610.
PR 18-JUN-1999; 99US-0139611.
PR 18-JUN-1999; 99US-0139612.
PR 18-JUN-1999; 99US-0139613.
PR 18-JUN-1999; 99US-0139614.
PR 18-JUN-1999; 99US-0139615.
PR 18-JUN-1999; 99US-0139616.
PR 18-JUN-1999; 99US-0139617.
PR 18-JUN-1999; 99US-0139618.
PR 18-JUN-1999; 99US-0139619.
PR 18-JUN-1999; 99US-0139620.
PR 18-JUN-1999; 99US-0139621.
PR 18-JUN-1999; 99US-0139622.
PR 18-JUN-1999; 99US-0139623.
PR 18-JUN-1999; 99US-0139624.
PR 18-JUN-1999; 99US-0139625.
PR 18-JUN-1999; 99US-0139626.
PR 18-JUN-1999; 99US-0139627.
PR 18-JUN-1999; 99US-0139628.
PR 18-JUN-1999; 99US-0139629.
PR 18-JUN-1999; 99US-0139630.
PR 18-JUN-1999; 99US-0139631.
PR 18-JUN-1999; 99US-0139632.
PR 18-JUN-1999; 99US-0139633.
PR 18-JUN-1999; 99US-0139634.
PR 18-JUN-1999; 99US-0139635.
PR 18-JUN-1999; 99US-0139636.
PR 18-JUN-1999; 99US-0139637.
PR 18-JUN-1999; 99US-0139638.
PR 18-JUN-1999; 99US-0139639.
PR 18-JUN-1999; 99US-0139640.
PR 18-JUN-1999; 99US-0139641.
PR 18-JUN-1999; 99US-0139642.
PR 18-JUN-1999; 99US-0139643.
PR 18-JUN-1999; 99US-0139644.
PR 18-JUN-1999; 99US-0139645.
PR 18-JUN-1999; 99US-0139646.
PR 18-JUN-1999; 99US-0139647.
PR 18-JUN-1999; 99US-0139648.
PR 18-JUN-1999; 99US-0139649.
PR 18-JUN-1999; 99US-0139650.
PR 18-JUN-1999; 99US-0139651.
PR 18-JUN-1999; 99US-0139652.
PR 18-JUN-1999; 99US-0139653.
PR 18-JUN-1999; 99US-0139654.
PR 18-JUN-1999; 99US-0139655.
PR 18-JUN-1999; 99US-0139656.
PR 18-JUN-1999; 99US-0139657.
PR 18-JUN-1999; 99US-0139658.
PR 18-JUN-1999; 99US-0139659.
PR 18-JUN-1999; 99US-0139660.
PR 18-JUN-1999; 99US-0139661.
PR 18-JUN-1999; 99US-0139662.
PR 18-JUN-1999; 99US-0139663.
PR 18-JUN-1999; 99US-0139664.
PR 18-JUN-1999; 99US-0139665.
PR 18-JUN-1999; 99US-0139666.
PR 18-JUN-1999; 99US-0139667.
PR 18-JUN-1999; 99US-0139668.
PR 18-JUN-1999; 99US-0139669.
PR 18-JUN-1999; 99US-0139670.
PR 18-JUN-1999; 99US-0139671.
PR 18-JUN-1999; 99US-0139672.
PR 18-JUN-1999; 99US-0139673.
PR 18-JUN-1999; 99US-0139674.
PR 18-JUN-1999; 99US-0139675.
PR 18-JUN-1999; 99US-0139676.
PR 18-JUN-1999; 99US-0139677.
PR 18-JUN-1999; 99US-0139678.
PR 18-JUN-1999; 99US-0139679.
PR 18-JUN-1999; 99US-0139680.
PR 18-JUN-1999; 99US-0139681.
PR 18-JUN-1999; 99US-0139682.
PR 18-JUN-1999; 99US-0139683.
PR 18-JUN-1999; 99US-0139684.
PR 18-JUN-1999; 99US-0139685.
PR 18-JUN-1999; 99US-0139686.
PR 18-JUN-1999; 99US-0139687.
PR 18-JUN-1999; 99US-0139688.
PR 18-JUN-1999; 99US-0139689.
PR 18-JUN-1999; 99US-0139690.
PR 18-JUN-1999; 99US-0139691.
PR 18-JUN-1999; 99US-0139692.
PR 18-JUN-1999; 99US-0139693.
PR 18-JUN-1999; 99US-0139694.
PR 18-JUN-1999; 99US-0139695.
PR 18-JUN-1999; 99US-0139696.
PR 18-JUN-1999; 99US-0139697.
PR 18-JUN-1999; 99US-0139698.
PR 18-JUN-1999; 99US-0139699.
PR 18-JUN-1999; 99US-0139700.
PR 18-JUN-1999; 99US-0139701.
PR 18-JUN-1999; 99US-0139702.
PR 18-JUN-1999; 99US-0139703.
PR 18-JUN-1999; 99US-0139704.
PR 18-JUN-1999; 99US-0139705.
PR 18-JUN-1999; 99US-0139706.
PR 18-JUN-1999; 99US-0139707.
PR 18-JUN-1999; 99US-0139708.
PR 18-JUN-1999; 99US-0139709.
PR 18-JUN-1999; 99US-0139710.
PR 18-JUN-1999; 99US-0139711.
PR 18-JUN-1999; 99US-0139712.
PR 18-JUN-1999; 99US-0139713.
PR 18-JUN-1999; 99US-0139714.
PR 18-JUN-1999; 99US-0139715.
PR 18-JUN-1999; 99US-0139716.
PR 18-JUN-1999; 99US-0139717.
PR 18-JUN-1999; 99US-0139718.
PR 18-JUN-1999; 99US-0139719.
PR 18-JUN-1999; 99US-0139720.
PR 18-JUN-1999; 99US-0139721.
PR 18-JUN-1999; 99US-0139722.
PR 18-JUN-1999; 99US-0139723.
PR 18-JUN-1999; 99US-0139724.
PR 18-JUN-1999; 99US-0139725.
PR 18-JUN-1999; 99US-0139726.
PR 18-JUN-1999; 99US-0139727.
PR 18-JUN-1999; 99US-0139728.
PR 18-JUN-1999; 99US-0139729.
PR 18-JUN-1999; 99US-0139730.
PR 18-JUN-1999; 99US-0139731.
PR 18-JUN-1999; 99US-0139732.
PR 18-JUN-1999; 99US-0139733.
PR 18-JUN-1999; 99US-0139734.
PR 18-JUN-1999; 99US-0139735.
PR 18-JUN-1999; 99US-0139736.
PR 18-JUN-1999; 99US-0139737.
PR 18-JUN-1999; 99US-0139738.
PR 18-JUN-1999; 99US-0139739.
PR 18-JUN-1999; 99US-0139740.
PR 18-JUN-1999; 99US-0139741.
PR 18-JUN-1999; 99US-0139742.
PR 18-JUN-1999; 99US-0139743.
PR 18-JUN-1999; 99US-0139744.
PR 18-JUN-1999; 99US-0139745.
PR 18-JUN-1999; 99US-0139746.
PR 18-JUN-1999; 99US-0139747.
PR 18-JUN-1999; 99US-0139748.
PR 18-JUN-1999; 99US-0139749.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139751.
PR 18-JUN-1999; 99US-0139752.
PR 18-JUN-1999; 99US-0139753.
PR 18-JUN-1999; 99US-0139754.
PR 18-JUN-1999; 99US-0139755.
PR 18-JUN-1999; 99US-0139756.
PR 18-JUN-1999; 99US-0139757.
PR 18-JUN-1999; 99US-0139758.
PR 18-JUN-1999; 99US-0139759.
PR 18-JUN-1999; 99US-0139760.
PR 18-JUN-1999; 99US-0139761.
PR 18-JUN-1999; 99US-0139762.
PR 18-JUN-1999; 99US-0139763.
PR 18-JUN-1999; 99US-0139764.
PR 18-JUN-1999; 99US-0139765.
PR 18-JUN-1999; 99US-0139766.
PR 18-JUN-1999; 99US-0139767.
PR 18-JUN-1999; 99US-0139768.
PR 18-JUN-1999; 99US-0139769.
PR 18-JUN-1999; 99US-0139770.
PR 18-JUN-1999; 99US-0139771.
PR 18-JUN-1999; 99US-0139772.
PR 18-JUN-1999; 99US-0139773.
PR 18-JUN-1999; 99US-0139774.
PR 18-JUN-1999; 99US-0139775.
PR 18-JUN-1999; 99US-0139776.
PR 18-JUN-1999; 99US-0139777.
PR 18-JUN-1999; 99US-0139778.
PR 18-JUN-1999; 99US-0139779.
PR 18-JUN-1999; 99US-0139780.
PR 18-JUN-1999; 99US-0139781.
PR 18-JUN-1999; 99US-0139782.
PR 18-JUN-1999; 99US-0139783.
PR 18-JUN-1999; 99US-0139784.
PR 18-JUN-1999; 99US-0139785.
PR 18-JUN-1999; 99US-0139786.
PR 18-JUN-1999; 99US-0139787.
PR 18-JUN-1999; 99US-0139788.
PR 18-JUN-1999; 99US-0139789.
PR 18-JUN-1999; 99US-0139790.
PR 18-JUN-1999; 99US-0139791.
PR 18-JUN-1999; 99US-0139792.
PR 18-JUN-1999; 99US-0139793.
PR 18-JUN-1999; 99US-0139794.
PR 18-JUN-1999; 99US-0139795.
PR 18-JUN-1999; 99US-0139796.
PR 18-JUN-1999; 99US-0139797.
PR 18-JUN-1999; 99US-0139798.
PR 18-JUN-1999; 99US-0139799.
PR 18-JUN-1999; 99US-0139800.
PR 18-JUN-1999; 99US-0139801.
PR 18-JUN-1999; 99US-0139802.
PR 18-JUN-1999; 99US-0139803.
PR 18-JUN-1999; 99US-0139804.
PR 18-JUN-1999; 99US-0139805.
PR 18-JUN-1999; 99US-0139806.
PR 18-JUN-1999; 99US-0139807.
PR 18-JUN-1999; 99US-0139808.
PR 18-JUN-1999; 99US-0139809.
PR 18-JUN-1999; 99US-0139810.
PR 18-JUN-1999; 99US-0139811.
PR 18-JUN-1999; 99US-0139812.
PR 18-JUN-1999; 99US-0139813.
PR 18-JUN-1999; 99US-0139814.
PR 18-JUN-1999; 99US-0139815.
PR 18-JUN-1999; 99US-0139816.
PR 18-JUN-1999; 99US-0139817.
PR 18-JUN-1999; 99US-0139818.
PR 18-JUN-1999; 99US-0139819.
PR 18-JUN-1999; 99US-0139820.
PR 18-JUN-1999; 99US-0139821.
PR 18-JUN-1999; 99US-0139822.
PR 18-JUN-1999; 99US-0139823.
PR 18-JUN-1999; 99US-0139824.
PR 18-JUN-1999; 99US-0139825.
PR 18-JUN-1999; 99US-0139826.
PR 18-JUN-1999; 99US-0139827.
PR 18-JUN-1999; 99US-0139828.
PR 18-JUN-1999; 99US-0139829.
PR 18-JUN-1999; 99US-0139830.
PR 18-JUN-1999; 99US-0139831.
PR 18-JUN-1999; 99US-0139832.
PR 18-JUN-1999; 99US-0139833.
PR 18-JUN-1999; 99US-0139834.
PR 18-JUN-1999; 99US-0139835.
PR 18-JUN-1999; 99US-01398
```





XX protein identification: signal transduction pathway: metabolic pathway:  
 KM hybridisation assay: genetic mapping, gene expression control, pre-mRNA  
 KM termination sequence  
 XX  
 US Arabidopsis thaliana.  
 FN EP1034405-A2.  
 XX  
 PD 06-SEP-2000.  
 XX  
 PE 25-FEB-2000: 2000EP-0301439.  
 XX  
 PR 25-FEB-1999: 99US-0121825.  
 PR 05-MAR-1999: 99US-012180.  
 PR 09-MAR-1999: 99US-0123548.  
 PR 23-MAR-1999: 99US-0123788.  
 PR 25-MAR-1999: 99US-0126264.  
 PR 29-MAR-1999: 99US-0126785.  
 PR 01-APR-1999: 99US-0127462.  
 PR 08-APR-1999: 99US-0128734.  
 PR 08-APR-1999: 99US-0128714.  
 PR 16-APR-1999: 99US-0129845.  
 PR 19-APR-1999: 99US-0130077.  
 PR 21-APR-1999: 99US-0130449.  
 PR 23-APR-1999: 99US-0130510.  
 PR 24-APR-1999: 99US-0130891.  
 PR 28-APR-1999: 99US-0131449.  
 PR 30-APR-1999: 99US-0132048.  
 PR 30-APR-1999: 99US-0132407.  
 PR 04-MAY-1999: 99US-0134884.  
 PR 05-MAY-1999: 99US-0132486.  
 PR 06-MAY-1999: 99US-0132486.  
 PR 06-MAY-1999: 99US-0132487.  
 PR 07-MAY-1999: 99US-0132863.  
 PR 11-MAY-1999: 99US-0134256.  
 PR 14-MAY-1999: 99US-0134218.  
 PR 14-MAY-1999: 99US-0134219.  
 PR 14-MAY-1999: 99US-0134221.  
 PR 14-MAY-1999: 99US-0134370.  
 PR 18-MAY-1999: 99US-0134768.  
 PR 19-MAY-1999: 99US-0134941.  
 PR 20-MAY-1999: 99US-0135124.  
 PR 21-MAY-1999: 99US-0135153.  
 PR 24-MAY-1999: 99US-0135629.  
 PR 25-MAY-1999: 99US-0136021.  
 PR 27-MAY-1999: 99US-0136392.  
 PR 28-MAY-1999: 99US-0136782.  
 PR 01-JUN-1999: 99US-0137222.  
 PR 03-JUN-1999: 99US-0137528.  
 PR 04-JUN-1999: 99US-0137502.  
 PR 07-JUN-1999: 99US-0137724.  
 PR 08-JUN-1999: 99US-0138094.  
 PR 10-JUN-1999: 99US-0138540.  
 PR 10-JUN-1999: 99US-0138847.  
 PR 14-JUN-1999: 99US-0139115.  
 PR 16-JUN-1999: 99US-0139452.  
 PR 16-JUN-1999: 99US-0139453.  
 PR 17-JUN-1999: 99US-0139492.  
 PR 18-JUN-1999: 99US-0139454.  
 PR 18-JUN-1999: 99US-0139460.  
 PR 18-JUN-1999: 99US-0139461.  
 PR 18-JUN-1999: 99US-0139462.  
 PR 18-JUN-1999: 99US-0139463.  
 PR 18-JUN-1999: 99US-0139463.  
 PR 18-JUN-1999: 99US-0139750.  
 PR 21-JUN-1999: 99US-0139763.  
 PR 22-JUN-1999: 99US-0139817.  
 PR 22-JUN-1999: 99US-0139899.

---

PR 23-JUN-1999: 99US-0140353.  
 PR 23-JUN-1999: 99US-0140354.  
 PR 24-JUN-1999: 99US-0140604.  
 PR 28-JUN-1999: 99US-0140874.  
 PR 29-JUN-1999: 99US-0140941.  
 PR 30-JUN-1999: 99US-0141287.  
 PR 01-JUL-1999: 99US-0141842.  
 PR 01-JUL-1999: 99US-0142154.  
 PR 02-JUL-1999: 99US-0142055.  
 PR 06-JUL-1999: 99US-0142390.  
 PR 08-JUL-1999: 99US-0142803.  
 PR 09-JUL-1999: 99US-0142920.  
 PR 12-JUL-1999: 99US-0142977.  
 PR 13-JUL-1999: 99US-0143542.  
 PR 14-JUL-1999: 99US-0143624.  
 PR 15-JUL-1999: 99US-0144005.  
 PR 16-JUL-1999: 99US-0144085.  
 PR 16-JUL-1999: 99US-0144086.  
 PR 16-JUL-1999: 99US-0144325.  
 PR 19-JUL-1999: 99US-0144331.  
 PR 19-JUL-1999: 99US-0144332.  
 PR 19-JUL-1999: 99US-0144333.  
 PR 19-JUL-1999: 99US-0144334.  
 PR 19-JUL-1999: 99US-0144335.  
 PR 20-JUL-1999: 99US-0144352.  
 PR 20-JUL-1999: 99US-0144632.  
 PR 20-JUL-1999: 99US-0144884.  
 PR 21-JUL-1999: 99US-0144814.  
 PR 21-JUL-1999: 99US-0145086.  
 PR 21-JUL-1999: 99US-0145088.  
 PR 22-JUL-1999: 99US-0145085.  
 PR 22-JUL-1999: 99US-0145087.  
 PR 22-JUL-1999: 99US-0145087.  
 PR 22-JUL-1999: 99US-0145089.  
 PR 22-JUL-1999: 99US-0145192.  
 PR 22-JUL-1999: 99US-0145145.  
 PR 23-JUL-1999: 99US-0145218.  
 PR 23-JUL-1999: 99US-0145224.  
 PR 26-JUL-1999: 99US-0145276.  
 PR 27-JUL-1999: 99US-0145913.  
 PR 27-JUL-1999: 99US-0145919.  
 PR 27-JUL-1999: 99US-0145919.  
 PR 28-JUL-1999: 99US-0145951.  
 PR 02-AUG-1999: 99US-0146386.  
 PR 02-AUG-1999: 99US-0146388.  
 PR 02-AUG-1999: 99US-0146389.  
 PR 03-AUG-1999: 99US-0147034.  
 PR 04-AUG-1999: 99US-0147204.  
 PR 04-AUG-1999: 99US-0147192.  
 PR 05-AUG-1999: 99US-0147260.  
 PR 05-AUG-1999: 99US-0147303.  
 PR 06-AUG-1999: 99US-0147416.  
 PR 09-AUG-1999: 99US-0147493.  
 PR 09-AUG-1999: 99US-0147935.  
 PR 10-AUG-1999: 99US-0148171.  
 PR 11-AUG-1999: 99US-0148319.  
 PR 12-AUG-1999: 99US-0148341.  
 PR 13-AUG-1999: 99US-0148565.  
 PR 13-AUG-1999: 99US-0148684.  
 PR 16-AUG-1999: 99US-0148688.  
 PR 17-AUG-1999: 99US-0149175.  
 PR 18-AUG-1999: 99US-0149426.  
 PR 20-AUG-1999: 99US-0149722.  
 PR 20-AUG-1999: 99US-0149723.  
 PR 20-AUG-1999: 99US-0149929.  
 PR 23-AUG-1999: 99US-0149932.  
 PR 23-AUG-1999: 99US-0149930.  
 PR 25-AUG-1999: 99US-0150566.  
 PR 26-AUG-1999: 99US-0150884.  
 PR 27-AUG-1999: 99US-0151065.  
 PR 27-AUG-1999: 99US-0151066.  
 PR 27-AUG-1999: 99US-0151080.  
 PR 30-AUG-1999: 99US-0151303.

```

PR 31-SEP-1999: 99US-0151438.
PR 01-SEP-1999: 99US-0151430.
PR 07-SEP-1999: 99US-0152363.
PR 10-SEP-1999: 99US-0153070.
PR 13-SEP-1999: 99US-0153758.
PR 15-SEP-1999: 99US-0154018.
PR 16-SEP-1999: 99US-0154039.
PR 20-SEP-1999: 99US-0154779.
PR 22-SEP-1999: 99US-0155139.
PR 23-SEP-1999: 99US-0155446.
PR 24-SEP-1999: 99US-0155659.
PR 28-SEP-1999: 99US-0156458.
PR 29-SEP-1999: 99US-0156596.
PR 04-OCT-1999: 99US-0157117.
PR 05-OCT-1999: 99US-0157763.
PR 06-OCT-1999: 99US-0157855.
PR 07-OCT-1999: 99US-0158029.
PR 08-OCT-1999: 99US-0158232.
PR 12-OCT-1999: 99US-0158369.
PR 13-OCT-1999: 99US-0159293.
PR 13-OCT-1999: 99US-0159294.
PR 13-OCT-1999: 99US-0159295.
PR 14-OCT-1999: 99US-0159330.
PR 14-OCT-1999: 99US-0159331.
PR 14-OCT-1999: 99US-0159637.
PR 14-OCT-1999: 99US-0159638.
PR 18-OCT-1999: 99US-0159584.
PR 21-OCT-1999: 99US-0150741.
PR 21-OCT-1999: 99US-0160767.
PR 21-OCT-1999: 99US-0160768.
PR 21-OCT-1999: 99US-0160770.
PR 21-OCT-1999: 99US-0160814.
PR 21-OCT-1999: 99US-0160815.
PR 22-OCT-1999: 99US-0160940.
PR 22-OCT-1999: 99US-0160941.
PR 22-OCT-1999: 99US-0160949.
PR 25-OCT-1999: 99US-0161404.
PR 25-OCT-1999: 99US-0161405.
PR 25-OCT-1999: 99US-0161406.
PR 26-OCT-1999: 99US-0161359.
PR 26-OCT-1999: 99US-0161360.
PR 26-OCT-1999: 99US-0161361.
PR 26-OCT-1999: 99US-0161361.
PR 28-OCT-1999: 99US-0161992.
PR 28-OCT-1999: 99US-0161993.
PR 29-OCT-1999: 99US-0162142.

```

Query Match 38.84: Score 716 5; PR 21: Length 389;  
 Best Local Similarity 39.34: Pred No 1 6e-64;  
 Matches 148; Conservative 72; Mismatches 130; Indels 27; Gaps 5;

```

OY 2 MNRGSSVYQNSSTQGVVSMAGPALKNVETPKKHILDAIMAILGQZACQINPTA 61
DB 9 MUKYGETSYKAKSLASGLISLIGRVMDALKKIMNSSEISIGIADIGSSQPNSL 68
OY 62 VISTIKRMKKKKEKLEINQOTLEQVYINDLPGNPNTLEFKGLSS--EVIGNKKEFV---- 115
DB 69 slmshvdlhlnpdlaprpelvtslndjpsndioyiaslpreifdrvmuukgylgr 128
OY 116 -----VYVMVPSGFGHGLTPKNSLHLVHSSYVHMLTPAKKLDLSEPLAL-----NKA 146
DB 129 gggscfsvaygslygrlfrtslthvssslhwlslypcreaeedrltadlenng 188
OY 167 KIVSKTSPPVAVKAYLSGHEHETMFLNARSGVAVNNGMVLILKQKSGSDPSLMSCH 246
DB 189 KIVSKTSPPKAKAYALGPTdlvfltrseelvgpgrmvslfgrtsldpttosecy 248
OY 227 TWELLAMIAELVUGLIDELDKLTFNIFSYFASLEVKDIVERDOSFTLDHTECFULD 285
DB 249 gwellaaglmamekajleekidafnapyaasaelkmvickegsfslidreispida 308
OY 286 -----SVEMDENKWKV-----GEKFTKVVKATLEPLISNUGPELIMDKLYDKTHIV 344

```

```

DB 309 eggisiseesypalatsrkpealasgrvssltiravcpmlpflgenvwdehlerakng 368
OY 335 SDLEAKLPKTSITLVL 351
DB 369 cyfyspprvalvlsl 385

RESULT 4
AAC22188
ID AAC22188 standard; Protein; 405 AA.
XX
AC AAC22188;
XX
ET 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment, SEQ ID NO: 25022.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
EP E033405-A2.
XX
PD 06-SEP-2000.
XX
PE 24-FEB-2000; 2000EP-0301439.
XX
PP 25-FEB-1999: 99US-0121825.
PP 05-MAR-1999: 99US-0123180.
PP 09-MAR-1999: 99US-0123548.
PP 23-MAR-1999: 99US-0125788.
PP 25-MAR-1999: 99US-0126264.
PP 29-MAR-1999: 99US-0126785.
PP 02-APR-1999: 99US-0127452.
PP 06-APR-1999: 99US-0128234.
PP 08-APR-1999: 99US-0128714.
PP 16-APR-1999: 99US-0129845.
PP 19-APR-1999: 99US-0130077.
PP 21-APR-1999: 99US-0130049.
PP 23-APR-1999: 99US-0130510.
PP 23-APR-1999: 99US-0130891.
PP 28-APR-1999: 99US-0131449.
PP 30-APR-1999: 99US-0132048.
PP 30-APR-1999: 99US-0132407.
PP 04-MAY-1999: 99US-0132484.
PP 05-MAY-1999: 99US-0132485.
PP 06-MAY-1999: 99US-0132485.
PP 06-MAY-1999: 99US-0132486.
PP 07-MAY-1999: 99US-0132487.
PP 07-MAY-1999: 99US-0132663.
PP 11-MAY-1999: 99US-0134256.
PP 14-MAY-1999: 99US-0134218.
PP 14-MAY-1999: 99US-0134219.
PP 14-MAY-1999: 99US-0134221.
PP 14-MAY-1999: 99US-0134370.
PP 18-MAY-1999: 99US-0134768.
PP 19-MAY-1999: 99US-0134941.
PP 20-MAY-1999: 99US-0135124.
PP 21-MAY-1999: 99US-0135353.
PP 24-MAY-1999: 99US-0135629.
PP 25-MAY-1999: 99US-0136021.
PP 27-MAY-1999: 99US-0136392.
PP 28-MAY-1999: 99US-0136782.
PP 02-JUN-1999: 99US-0137252.
PP 03-JUN-1999: 99US-0137528.
PP 04-JUN-1999: 99US-0137502.
PP 07-JUN-1999: 99US-0137724.
PP 08-JUN-1999: 99US-0138094.
PP 10-JUN-1999: 99US-0138540.
PP 10-JUN-1999: 99US-0138847.
PP 14-JUN-1999: 99US-0139119.

```

```

PR 16-JUN-1999; 990S-0139452.
PR 16-JUN-1999; 990S-0139453.
PR 17-JUN-1999; 990S-0139492.
PR 18-JUN-1999; 990S-0139454.
PR 18-JUN-1999; 990S-0139455.
PR 18-JUN-1999; 990S-0139456.
PR 18-JUN-1999; 990S-0139457.
PR 18-JUN-1999; 990S-0139458.
PR 18-JUN-1999; 990S-0139459.
PR 18-JUN-1999; 990S-0139460.
PR 18-JUN-1999; 990S-0139461.
PR 18-JUN-1999; 990S-0139462.
PR 18-JUN-1999; 990S-0139463.
PR 18-JUN-1999; 990S-0139750.
PR 18-JUN-1999; 990S-0139753.
PR 21-JUN-1999; 990S-0139817.
PR 22-JUN-1999; 990S-0139819.
PR 23-JUN-1999; 990S-0140353.
PR 24-JUN-1999; 990S-0140354.
PR 24-JUN-1999; 990S-0140695.
PR 28-JUN-1999; 990S-0140823.
PR 29-JUN-1999; 990S-0140991.
PR 30-JUN-1999; 990S-0141287.
PR 01-JUL-1999; 990S-0141842.
PR 01-JUL-1999; 990S-0142154.
PR 02-JUL-1999; 990S-0142055.
PR 06-JUL-1999; 990S-0142360.
PR 08-JUL-1999; 990S-0142803.
PR 09-JUL-1999; 990S-0142920.
PR 12-JUL-1999; 990S-0142977.
PR 13-JUL-1999; 990S-014342.
PR 14-JUL-1999; 990S-0143624.
PR 15-JUL-1999; 990S-0144005.
PR 16-JUL-1999; 990S-0144085.
PR 16-JUL-1999; 990S-0144086.
PR 19-JUL-1999; 990S-0144425.
PR 19-JUL-1999; 990S-0144431.
PR 19-JUL-1999; 990S-0144332.
PR 19-JUL-1999; 990S-0144333.
PR 19-JUL-1999; 990S-0144334.
PR 19-JUL-1999; 990S-0144335.
PR 20-JUL-1999; 990S-0144352.
PR 20-JUL-1999; 990S-0144632.
PR 20-JUL-1999; 990S-0144884.
PR 21-JUL-1999; 990S-0144814.
PR 21-JUL-1999; 990S-0145086.
PR 21-JUL-1999; 990S-0145088.
PR 22-JUL-1999; 990S-0145085.
PR 22-JUL-1999; 990S-0145087.
PR 22-JUL-1999; 990S-0145089.
PR 22-JUL-1999; 990S-0145192.
PR 23-JUL-1999; 990S-0145145.
PR 23-JUL-1999; 990S-0145218.
PR 23-JUL-1999; 990S-0145224.
PR 25-JUL-1999; 990S-0145276.
PR 27-JUL-1999; 990S-0145913.
PR 27-JUL-1999; 990S-0145918.
PR 27-JUL-1999; 990S-0145919.
PR 28-JUL-1999; 990S-0145951.
PR 02-AUG-1999; 990S-0146386.
PR 02-AUG-1999; 990S-0146388.
PR 02-AUG-1999; 990S-0146389.
PR 03-AUG-1999; 990S-0147038.
PR 04-AUG-1999; 990S-0147204.
PR 04-AUG-1999; 990S-0147302.
PR 05-AUG-1999; 990S-0147192.
PR 05-AUG-1999; 990S-0147260.
PR 06-AUG-1999; 990S-0147303.
PR 06-AUG-1999; 990S-0147416.
PR 09-AUG-1999; 990S-0147493.
PR 09-AUG-1999; 990S-0147935.
PR 10-AUG-1999; 990S-0148171.
PR 11-AUG-1999; 990S-0148319.

```

```

PR 12-AUG-1999; 990S-0148341.
PR 13-AUG-1999; 990S-0148365.
PR 13-AUG-1999; 990S-0148684.
PR 16-AUG-1999; 990S-0149368.
PR 17-AUG-1999; 990S-0149175.
PR 18-AUG-1999; 990S-0149426.
PR 20-AUG-1999; 990S-0149722.
PR 20-AUG-1999; 990S-0149723.
PR 20-AUG-1999; 990S-0149929.
PR 23-AUG-1999; 990S-0149929.
PR 23-AUG-1999; 990S-0149930.
PR 25-AUG-1999; 990S-0150566.
PR 26-AUG-1999; 990S-0150884.
PR 27-AUG-1999; 990S-0151065.
PR 27-AUG-1999; 990S-0151066.
PR 27-AUG-1999; 990S-0151080.
PR 30-AUG-1999; 990S-0151303.
PR 31-AUG-1999; 990S-0151438.
PR 01-SEP-1999; 990S-0151930.
PR 07-SEP-1999; 990S-0152363.
PR 10-SEP-1999; 990S-0153070.
PR 13-SEP-1999; 990S-0153758.
PR 15-SEP-1999; 990S-0154018.
PR 16-SEP-1999; 990S-0154039.
PR 20-SEP-1999; 990S-0154779.
PR 22-SEP-1999; 990S-0155139.
PR 23-SEP-1999; 990S-0155486.
PR 24-SEP-1999; 990S-0155659.
PR 28-SEP-1999; 990S-0156458.
PR 29-SEP-1999; 990S-0156596.
PR 04-OCT-1999; 990S-0157117.
PR 05-OCT-1999; 990S-0157753.
PR 06-OCT-1999; 990S-0157865.
PR 07-OCT-1999; 990S-0158029.
PR 08-OCT-1999; 990S-0158232.
PR 12-OCT-1999; 990S-0158369.
PR 13-OCT-1999; 990S-0159293.
PR 13-OCT-1999; 990S-0159294.
PR 14-OCT-1999; 990S-0159329.
PR 14-OCT-1999; 990S-0159330.
PR 14-OCT-1999; 990S-0159331.
PR 14-OCT-1999; 990S-0159637.
PR 14-OCT-1999; 990S-0159638.
PR 18-OCT-1999; 990S-0159584.
PR 21-OCT-1999; 990S-0160741.
PR 21-OCT-1999; 990S-0160767.
PR 21-OCT-1999; 990S-0160768.
PR 21-OCT-1999; 990S-0160770.
PR 21-OCT-1999; 990S-0160814.
PR 21-OCT-1999; 990S-0160815.
PR 22-OCT-1999; 990S-0160980.
PR 22-OCT-1999; 990S-0160981.
PR 22-OCT-1999; 990S-0160989.
PR 23-OCT-1999; 990S-0161404.
PR 25-OCT-1999; 990S-0161405.
PR 25-OCT-1999; 990S-0161406.
PR 26-OCT-1999; 990S-0161359.
PR 26-OCT-1999; 990S-0161360.
PR 26-OCT-1999; 990S-0161361.
PR 28-OCT-1999; 990S-0161920.
PR 28-OCT-1999; 990S-0161992.
PR 28-OCT-1999; 990S-0161993.
PR 29-OCT-1999; 990S-0162142.

```

## Query Match

38.8%; Score 716.5; DB 21; Length 405;

Best Local Similarity 39.8%; Pref No 1.7e+63;

Matches 148; Conservative 72; Mismatches 140; Indels 27; Gaps 5;

```

QY 2 MNRGCPSSYAQNSPTQGVASMAQPAIFNAVPIIFSRDPIHQAIAAALIGCAAGPNTFA 61
  ||| ||||| ||| . . . . . : ||| : ||||| ||| :
2b 25 mdkkgpctcyahstqagctststlfrtwcdaklkhlmshvctststgjdldqstpsll 84

```



```

PR 27-JUL-1999; 990S-0145618.
PR 27-JUL-1999; 990S-0145619.
PR 28-JUL-1999; 990S-0145951.
PR 02-AUG-1999; 990S-0146386.
PR 02-AUG-1999; 990S-0146388.
PR 02-AUG-1999; 990S-0146389.
PR 03-AUG-1999; 990S-0147048.
PR 04-AUG-1999; 990S-0147204.
PR 04-AUG-1999; 990S-0147302.
PR 05-AUG-1999; 990S-0147792.
PR 05-AUG-1999; 990S-0147260.
PR 06-AUG-1999; 990S-0147303.
PR 06-AUG-1999; 990S-0147416.
PR 09-AUG-1999; 990S-0147493.
PR 09-AUG-1999; 990S-0147935.
PR 10-AUG-1999; 990S-0148171.
PR 11-AUG-1999; 990S-0148319.
PR 12-AUG-1999; 990S-0148341.
PR 13-AUG-1999; 990S-0148565.
PR 16-AUG-1999; 990S-0148684.
PR 17-AUG-1999; 990S-0149175.
PR 18-AUG-1999; 990S-0149426.
PR 20-AUG-1999; 990S-0149723.
PR 20-AUG-1999; 990S-0149929.
PR 23-AUG-1999; 990S-0149902.
PR 24-AUG-1999; 990S-0149940.
PR 25-AUG-1999; 990S-0150566.
PR 26-AUG-1999; 990S-0150884.
PR 27-AUG-1999; 990S-0151065.
PR 27-AUG-1999; 990S-0151066.
PR 27-AUG-1999; 990S-0151080.
PR 31-AUG-1999; 990S-0151303.
PR 01-SEP-1999; 990S-0151438.
PR 07-SEP-1999; 990S-0151930.
PR 10-SEP-1999; 990S-0152363.
PR 13-SEP-1999; 990S-0153070.
PR 15-SEP-1999; 990S-0153758.
PR 16-SEP-1999; 990S-0154018.
PR 20-SEP-1999; 990S-0154039.
PR 22-SEP-1999; 990S-0154779.
PR 24-SEP-1999; 990S-0155139.
PR 24-SEP-1999; 990S-0155486.
PR 28-SEP-1999; 990S-0155659.
PR 29-SEP-1999; 990S-0156458.
PR 04-OCT-1999; 990S-0157117.
PR 05-OCT-1999; 990S-0157753.
PR 06-OCT-1999; 990S-0157865.
PR 07-OCT-1999; 990S-0158029.
PR 08-OCT-1999; 990S-0158232.
PR 12-OCT-1999; 990S-0158364.
PR 13-OCT-1999; 990S-0159293.
PR 13-OCT-1999; 990S-0159294.
PR 13-OCT-1999; 990S-0159295.
PR 14-OCT-1999; 990S-0159329.
PR 14-OCT-1999; 990S-0159330.
PR 14-OCT-1999; 990S-0159431.
PR 14-OCT-1999; 990S-0159637.
PR 14-OCT-1999; 990S-0159638.
PR 18-OCT-1999; 990S-0159584.
PR 21-OCT-1999; 990S-0160741.
PR 21-OCT-1999; 990S-0160767.
PR 21-OCT-1999; 990S-0160768.
PR 21-OCT-1999; 990S-0160771.
PR 21-OCT-1999; 990S-0160814.
PR 21-OCT-1999; 990S-0160815.
PR 22-OCT-1999; 990S-0160880.
PR 22-OCT-1999; 990S-0160981.
PR 22-OCT-1999; 990S-0160989.
PR 25-OCT-1999; 990S-0161404.
PR 25-OCT-1999; 990S-0161405.

```

```

PR 25-OCT-1999; 990S-0161406.
PR 26-OCT-1999; 990S-0161359.
PR 26-OCT-1999; 990S-0161360.
PR 26-OCT-1999; 990S-0161361.
PR 28-OCT-1999; 990S-0161920.
PR 28-OCT-1999; 990S-0161922.
PR 28-OCT-1999; 990S-0161992.
PR 29-OCT-1999; 990S-0162342.
PR 29-OCT-1999; 990S-0162342.

Query Match          37.2% Score 687; DB 21; Length 349;
Best Local Similarity 40.2% Fred. No 1 3e-60;
Matches 144, Conservative 70, Mismatches 128, Indels 16, Gaps 6;

QY 2 MNRGFFSSAAGNSFFGVASMAAGAI-ENAVETTFSPDREHLGAINAHGCAAGPRTF 60
Db 1 msqgddqansysansrlqkxvismakpvlvrrlccommldlply-1kvaolqssqgnsl 59
QY 61 AVISTIKRMMEKKCPREINCOLEILOYVNLDFGNDFNLKCL---SSEVIGNCEEVP 116
Db 60 latfeintindvncgvnknspedlccclndfndntlntkfyptlncmlndss--- 116
QY 117 CYWKGVPSTFGRLEPPNSLIHLWSSSYVMWLTQAPKGLTSEGLALNKGIYISKTSPP 176
Db 117 cygqpgstfyslftstnlhltsygrtlwlskcy- ch-nd-nllytlsssf- 170
QY 177 VVREAVYISOPHPPTMPLNARSGFVVPNGCMVLLRGHO-CSDPDMSQCTWELLAMAI 235
Db 171 saykayinqfkdntmlrlfseerlsrgawltltgtrnlndpplrydchhwtllsssl 230
Db 231 rdlvfyglvsqsklntmlrlfseerlsrgawltltgtrnlndpplrydchhwtllsssf 290
QY 296 VPGKEFTKVVAPTEPPISNPGPELMDKTYDKPTTHIVSDLEALPRTTSLLVLSK 353
Db 291 cagrnasnglravsepmllshlqgeellcltkyayhvcuqancnktcvsivsltk 348

RESULT 6
AAG37440
ID AAG37440 standard; protein; 349 AA.
XX
AC AAG37440;
XX
L1 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SRQ ID NO: 46036.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
FF 25 FEB 2000; J060EP-0001439.
XX
PR 25-FEB-1999; 990S-0123825.
PR 05-MAR-1999; 990S-0123180.
PR 09-MAR-1999; 990S-0123548.
PR 23-MAR-1999; 990S-0125788.
PR 23-MAR-1999; 990S-0125785.
PR 01-APR-1999; 990S-0123462.
PR 06-APR-1999; 990S-0128234.
PR 08-APR-1999; 990S-0128714.
PR 16-APR-1999; 990S-0129845.
PR 19-APR-1999; 990S-0130077.
PR 21-APR-1999; 990S-0130449.
PR 23-APR-1999; 990S-0130510.

```







```

PR 13-JUL-1999: 99US-0143642
PR 14-JUL-1999: 99US-0143624
PR 15-JUL-1999: 99US-0144005
PR 16-JUL-1999: 99US-0144085
PR 16-JUL-1999: 99US-0144086
PR 19-JUL-1999: 99US-0144325
PR 19-JUL-1999: 99US-0144331
PR 19-JUL-1999: 99US-0144332
PR 19-JUL-1999: 99US-0144334
PR 19-JUL-1999: 99US-0144335
PR 20-JUL-1999: 99US-0144352
PR 20-JUL-1999: 99US-0144684
PR 21-JUL-1999: 99US-0144814
PR 21-JUL-1999: 99US-0145086
PR 21-JUL-1999: 99US-0145088
PR 22-JUL-1999: 99US-0145085
PR 22-JUL-1999: 99US-0145087
PR 22-JUL-1999: 99US-0145089
PR 22-JUL-1999: 99US-0145192
PR 23-JUL-1999: 99US-0145145
PR 23-JUL-1999: 99US-0145218
PR 23-JUL-1999: 99US-0145224
PR 26-JUL-1999: 99US-0145276
PR 27-JUL-1999: 99US-0145913
PR 27-JUL-1999: 99US-0145918
PR 27-JUL-1999: 99US-0145919
PR 28-JUL-1999: 99US-0145951
PR 02-AUG-1999: 99US-0146389
PR 02-AUG-1999: 99US-0146388
PR 03-AUG-1999: 99US-0147036
PR 04-AUG-1999: 99US-0147204
PR 04-AUG-1999: 99US-0147302
PR 05-AUG-1999: 99US-0147192
PR 05-AUG-1999: 99US-0147260
PR 06-AUG-1999: 99US-0147203
PR 06-AUG-1999: 99US-0147416
PR 09-AUG-1999: 99US-0147493
PR 09-AUG-1999: 99US-0147935
PR 10-AUG-1999: 99US-0148171
PR 11-AUG-1999: 99US-0148319
PR 12-AUG-1999: 99US-0148341
PR 13-AUG-1999: 99US-0148565
PR 13-AUG-1999: 99US-0148684
PR 16-AUG-1999: 99US-0149368
PR 17-AUG-1999: 99US-0149175
PR 18-AUG-1999: 99US-0149426
PR 20-AUG-1999: 99US-0149722
PR 20-AUG-1999: 99US-0149723
PR 20-AUG-1999: 99US-0149829
PR 23-AUG-1999: 99US-0149902
PR 23-AUG-1999: 99US-0149939
PR 25-AUG-1999: 99US-0150566
PR 26-AUG-1999: 99US-0150884
PR 27-AUG-1999: 99US-0151065
PR 27-AUG-1999: 99US-0151065
PR 27-AUG-1999: 99US-0151080
PR 30-AUG-1999: 99US-0151308
PR 31-AUG-1999: 99US-0151338
PR 01-SEP-1999: 99US-0151430
PR 07-SEP-1999: 99US-0152363
PR 10-SEP-1999: 99US-0153070
PR 13-SEP-1999: 99US-0153758
PR 15-SEP-1999: 99US-0154018
PR 16-SEP-1999: 99US-0154039
PR 20-SEP-1999: 99US-0154779
PR 22-SEP-1999: 99US-0155139
PR 23-SEP-1999: 99US-0155486
PR 24-SEP-1999: 99US-0155559
PR 28-SEP-1999: 99US-0156458
PR 29-SEP-1999: 99US-0156596

```

```

PR 04-OCT-1999: 99US-0157117
PR 05-OCT-1999: 99US-0157753
PR 06-OCT-1999: 99US-0157865
PR 07-OCT-1999: 99US-0158029
PR 08-OCT-1999: 99US-0158232
PR 12-OCT-1999: 99US-0158459
PR 13-OCT-1999: 99US-0159293
PR 13-OCT-1999: 99US-0159294
PR 13-OCT-1999: 99US-0159295
PR 14-OCT-1999: 99US-0159329
PR 14-OCT-1999: 99US-0159330
PR 14-OCT-1999: 99US-0159331
PR 14-OCT-1999: 99US-0159637
PR 14-OCT-1999: 99US-0159638
PR 18-OCT-1999: 99US-0159584
PR 21-OCT-1999: 99US-0160741
PR 21-OCT-1999: 99US-0160741
PR 21-OCT-1999: 99US-0160767
PR 21-OCT-1999: 99US-0160768
PR 21-OCT-1999: 99US-0160770
PR 21-OCT-1999: 99US-0160814
PR 21-OCT-1999: 99US-0160815
PR 22-OCT-1999: 99US-0160980
PR 22-OCT-1999: 99US-0160981
PR 23-OCT-1999: 99US-0160989
PR 25-OCT-1999: 99US-0161404
PR 25-OCT-1999: 99US-0161405
PR 26-OCT-1999: 99US-0161406
PR 26-OCT-1999: 99US-0161409
PR 26-OCT-1999: 99US-0161360
PR 26-OCT-1999: 99US-0161361
PR 26-OCT-1999: 99US-0161920
PR 28-OCT-1999: 99US-0161993
PR 28-OCT-1999: 99US-0161993
PR 29-OCT-1999: 99US-0162147

Query Match: 37.2% Score 687: DB 21: Length 379:
Post local Similarity 40.2% Pred No 15a-60:
Matches 144: Conservative 70: Mismatches 128: Indels 16: Gaps 6:

QY 2 MNQGEKSSYAQNSPFGVAVASMAQAL-ENAVEILLESKPHLQALMAADQAGAPNIF 60
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 31 msygdgansysansrlqkxvlsmakpvlvrnlocommldlfty-ikvaqlgssgnstf 89

QY 61 AVISTIKPMKEKKCPFINQOTLEIYVINDIFGNPENTIFKCI-----SPFVIGNCEVP 116
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 90 lafelintlnvleqnvkospeldecindpendnttktvptfinkelmtnkss--- 146

QY 117 CYVMVSPGFHGRKLEPNSLILVHSSSVHWLQAQKSLDSKGLNKGKLYISKISPP 176
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 147 cftvgagstysltstfosvltshsgytlhwlskvp-----eklenudgldyltsssfq 200

QY 177 VVEKAVYSGEHPPTMNAKSGEVVNRGMVLLIRPQ-CEQVNSMSTPTWPLAMAT 235
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 201 svghyvtqtkltmcltssvssgmavcltfgjodldpdyldcclaklssfcl 260

QY 236 AEIVSGILDEEKLDTFNIPSYFASLEEVKLEVEFGSPTIHIGFPLSVEMGEKMK 295
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 261 rdlvtelgvscldatmmpfygnvqelkevqgkcslelnelshtyldlytyceddt 320

QY 296 VPGKPKTKVPAETPTIISNGSPETIMKLYNKPFIHIVSULEAKLPKTSIIYLAK 353
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 323 gaptgaangltvssqgmclahfgcecltcltkyayhvqhancnktlysvlssltk 378

RESULT 8
AAC37439
ID AAC37439 standard: Protein: 379 AA.
AC AAC37439:
XX
XX 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 46035.

```

XX protein identification, signal transduction pathway; metabolism pathway;  
KM hybridization assay, genetic mapping, gene expression control; promoter;  
KW termination sequence.  
XX Arabidopsis thaliana.  
XX EP1033405-A2.  
XX 06-SEP-2000.  
XX 25-FEB-2000; 2000EP-0301439  
XX  
PE 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127452.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135241.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137602.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139917.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 24-JUN-1999; 99US-0140354.  
PR 28-JUN-1999; 99US-0140695.  
PR 29-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140941.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142965.  
PR 06-JUL-1999; 99US-0142969.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144432.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 23-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145376.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147294.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147392.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 11-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148665.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149924.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149926.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.

```

PR 31-AUG-1999; 990S-0151438.
PR 01-SEP-1999; 990S-0151430.
PR 07-SEP-1999; 990S-0152363.
PR 10-SEP-1999; 990S-0153070.
PR 13-SEP-1999; 990S-0153758.
PR 15-SEP-1999; 990S-0154018.
PR 16-SEP-1999; 990S-0154036.
PR 20-SEP-1999; 990S-0154779.
PR 22-SEP-1999; 990S-0155139.
PR 23-SEP-1999; 990S-0155486.
PR 24-SEP-1999; 990S-0155559.
PR 28-SEP-1999; 990S-0156458.
PR 29-SEP-1999; 990S-0156596.
PR 04-OCT-1999; 990S-0157117.
PR 05-OCT-1999; 990S-0157753.
PR 06-OCT-1999; 990S-0157865.
PR 07-OCT-1999; 990S-0158029.
PR 08-OCT-1999; 990S-0158232.
PR 12-OCT-1999; 990S-0158369.
PR 13-OCT-1999; 990S-0159293.
PR 13-OCT-1999; 990S-0159294.
PR 13-OCT-1999; 990S-0159295.
PR 14-OCT-1999; 990S-0159429.
PR 14-OCT-1999; 990S-0159930.
PR 14-OCT-1999; 990S-0159937.
PR 14-OCT-1999; 990S-0159937.
PR 18-OCT-1999; 990S-0159584.
PR 18-OCT-1999; 990S-0159584.
PR 21-OCT-1999; 990S-0160741.
PR 21-OCT-1999; 990S-0160767.
PR 21-OCT-1999; 990S-0160768.
PR 21-OCT-1999; 990S-0160770.
PR 21-OCT-1999; 990S-0160814.
PR 21-OCT-1999; 990S-0160815.
PR 22-OCT-1999; 990S-0160980.
PR 22-OCT-1999; 990S-0160981.
PR 22-OCT-1999; 990S-0160989.
PR 25-OCT-1999; 990S-0161404.
PR 25-OCT-1999; 990S-0161405.
PR 25-OCT-1999; 990S-0161406.
PR 26-OCT-1999; 990S-0161450.
PR 26-OCT-1999; 990S-0161450.
PR 26-OCT-1999; 990S-0161361.
PR 26-OCT-1999; 990S-0161361.
PR 28-OCT-1999; 990S-0161992.
PR 28-OCT-1999; 990S-0161993.
PR 29-OCT-1999; 990S-0162142.

```

```

Query Map: 37 38 Score 687 106 21 Length 473
Best Local Similarity 40 28; Prod No 1 5e-60;
Matches 144; Conservative 70; Mismatches 128; Indels 16; Gaps 6;

```

```

CY 2 MNRKQSSSTAVNSSTLVVAVSMAGPAL-ENAVVILSKDFHQAIAADLQAGAPNLF 60
    . . . . . | . . . . . | . . . . . | . . . . . | . . . . . |
DB 31 msygydasyasasr ldkkylsmakvlylvnteeamndldply-lkvelgssyqsf 89
    . . . . . | . . . . . | . . . . . | . . . . . | . . . . . |
CY 61 AVISTIRMMKKRPRENCCFLVYVNLNLENDNFILKEL---SSVGVNKEEVP 119
    . . . . . | . . . . . | . . . . . | . . . . . | . . . . . |
DB 90 lafeilnltlvtoqhnkspcidvrlodipodfntctfkpfvfnkeimtnkss--- 149
    . . . . . | . . . . . | . . . . . | . . . . . | . . . . . |
CY 117 CYVMVPSSEHLEPNSEH VHSYSVHMLTGAHKLSPRIAI NQKLVISKSTP 179
    . . . . . | . . . . . | . . . . . | . . . . . | . . . . . |
DB 147 cfvygafgsfysrlfssstlhlhssyalhlakylf---klomknkgnlyrfsespa 200
    . . . . . | . . . . . | . . . . . | . . . . . | . . . . . |
CY 177 VVRKATLSQHEFTMLNAKSDEVVNDIMVILKRPQ--SNPSIMWGSFTWELLAMAI 245
    . . . . . | . . . . . | . . . . . | . . . . . | . . . . . |
DB 201 saykaylnqtkdftmlrlfscvfvsngrvylfignrlndpilyrccbfwtllnsel 260
    . . . . . | . . . . . | . . . . . | . . . . . | . . . . . |
CY 236 AELVSGGLIDDKLDITNIPSYASLEEKDIVERGDSFTIDIEGFDSDVPMQNDKW 295
    . . . . . | . . . . . | . . . . . | . . . . . | . . . . . |
DB 261 rdltfeglvyskldatfmpfydgnvqelkevlgkcsfclnecstgldlghyfyedcd 320
    . . . . . | . . . . . | . . . . . | . . . . . | . . . . . |
CY 296 VRGRFKTKVRAPEPIISNQPPEFIMDKLYRFTHLVSDLEAKLPKTSILLVISK 353
    . . . . . | . . . . . | . . . . . | . . . . . | . . . . . |

```

```

DB 321 caagmcaagtrawsepmllahgeclldldkkyahvghacnroklwslvstlk 378
RESULT 9
ID AAG12688 standard; Protein: 327 AA.
XX AAG12688.
AC AAG12688.
DE 17-OCT-2000 (first entry)
XX
XX Arabidopsis thaliana protein fragment SEQ ID No: 11899.
XX protein identification: signal transduction pathway; metabolic pathway;
XX hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX Arabidopsis thaliana.
XX EPI031405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-1999; 990S-0121825.
XX 05-MAY-1999; 990S-0123180.
XX 03-MAR-1999; 990S-0123548.
XX 23-MAR-1999; 990S-0125788.
XX 25-MAR-1999; 990S-0125788.
XX 29-MAR-1999; 990S-0126785.
XX 01-APR-1999; 990S-0127462.
XX 06-APR-1999; 990S-0128234.
XX 08-APR-1999; 990S-0128714.
XX 16-APR-1999; 990S-0129845.
XX 19-APR-1999; 990S-0130077.
XX 21-APR-1999; 990S-0130449.
XX 23-APR-1999; 990S-0130510.
XX 23-APR-1999; 990S-0130891.
XX 28-APR-1999; 990S-0131449.
XX 30-APR-1999; 990S-0132048.
XX 30-APR-1999; 990S-0132407.
XX 04-MAY-1999; 990S-0132484.
XX 05-MAY-1999; 990S-0132485.
XX 06-MAY-1999; 990S-0132486.
XX 06-MAY-1999; 990S-0132487.
XX 07-MAY-1999; 990S-0132863.
XX 11-MAY-1999; 990S-0134256.
XX 14-MAY-1999; 990S-0134218.
XX 14-MAY-1999; 990S-0134219.
XX 14-MAY-1999; 990S-0134221.
XX 14-MAY-1999; 990S-0134370.
XX 18-MAY-1999; 990S-0134768.
XX 19-MAY-1999; 990S-0134941.
XX 20-MAY-1999; 990S-0135124.
XX 21-MAY-1999; 990S-0135353.
XX 24-MAY-1999; 990S-0135629.
XX 25-MAY-1999; 990S-0136021.
XX 27-MAY-1999; 990S-0136392.
XX 28-MAY-1999; 990S-0136782.
XX 01-JUN-1999; 990S-0137222.
XX 01-JUN-1999; 990S-0137228.
XX 04-JUN-1999; 990S-0137502.
XX 07-JUN-1999; 990S-0137724.
XX 08-JUN-1999; 990S-0138094.
XX 10-JUN-1999; 990S-0138450.
XX 10-JUN-1999; 990S-0138847.
XX 14-JUN-1999; 990S-0139119.
XX 16-JUN-1999; 990S-0139452.
XX 16-JUN-1999; 990S-0139453.
XX 17-JUN-1999; 990S-0139492.
XX 18-JUN-1999; 990S-0139454.

```







PR 07-MAY-1999: 99US-0132863  
PR 11-MAY-1999: 99US-0134256  
PR 14-MAY-1999: 99US-0134218  
PR 14-MAY-1999: 99US-0134219  
PR 14-MAY-1999: 99US-0134221  
PR 14-MAY-1999: 99US-0134370  
PR 18-MAY-1999: 99US-0134768  
PR 19-MAY-1999: 99US-0134941  
PR 20-MAY-1999: 99US-0135124  
PR 21-MAY-1999: 99US-0135353  
PR 24-MAY-1999: 99US-0135529  
PR 25-MAY-1999: 99US-0136021  
PR 27-MAY-1999: 99US-0136392  
PR 28-MAY-1999: 99US-0137232  
PR 01-JUN-1999: 99US-0137528  
PR 03-JUN-1999: 99US-0137528  
PR 04-JUN-1999: 99US-0137502  
PR 07-JUN-1999: 99US-0137724  
PR 08-JUN-1999: 99US-0138094  
PR 10-JUN-1999: 99US-0138540  
PR 10-JUN-1999: 99US-0138847  
PR 14-JUN-1999: 99US-0139119  
PR 16-JUN-1999: 99US-0139452  
PR 16-JUN-1999: 99US-0139453  
PR 17-JUN-1999: 99US-0139492  
PR 18-JUN-1999: 99US-0139454  
PR 18-JUN-1999: 99US-0139454  
PR 18-JUN-1999: 99US-0139456  
PR 18-JUN-1999: 99US-0139457  
PR 18-JUN-1999: 99US-0139458  
PR 18-JUN-1999: 99US-0139459  
PR 18-JUN-1999: 99US-0139460  
PR 18-JUN-1999: 99US-0139461  
PR 18-JUN-1999: 99US-0139462  
PR 18-JUN-1999: 99US-0139463  
PR 18-JUN-1999: 99US-0139750  
PR 18-JUN-1999: 99US-0139764  
PR 21-JUN-1999: 99US-0139817  
PR 22-JUN-1999: 99US-0139899  
PR 23-JUN-1999: 99US-0140354  
PR 23-JUN-1999: 99US-0140354  
PR 24-JUN-1999: 99US-0140695  
PR 26-JUN-1999: 99US-0140823  
PR 29-JUN-1999: 99US-0140991  
PR 30-JUN-1999: 99US-0141287  
PR 01-JUL-1999: 99US-0141842  
PR 01-JUL-1999: 99US-0142184  
PR 02-JUL-1999: 99US-0142356  
PR 06-JUL-1999: 99US-0142390  
PR 08-JUL-1999: 99US-0142403  
PR 09-JUL-1999: 99US-0142401  
PR 12-JUL-1999: 99US-0142977  
PR 14-JUL-1999: 99US-0143542  
PR 14-JUL-1999: 99US-0143624  
PR 15-JUL-1999: 99US-0144005  
PR 16-JUL-1999: 99US-0144085  
PR 16-JUL-1999: 99US-0144086  
PR 19-JUL-1999: 99US-0144325  
PR 19-JUL-1999: 99US-0144331  
PR 19-JUL-1999: 99US-0144332  
PR 19-JUL-1999: 99US-0144334  
PR 19-JUL-1999: 99US-0144334  
PR 19-JUL-1999: 99US-0144335  
PR 20-JUL-1999: 99US-0144352  
PR 20-JUL-1999: 99US-0144362  
PR 20-JUL-1999: 99US-0144364  
PR 21-JUL-1999: 99US-0144414  
PR 21-JUL-1999: 99US-0145086  
PR 21-JUL-1999: 99US-0145088  
PR 22-JUL-1999: 99US-0145085  
PR 22-JUL-1999: 99US-0145087  
PR 22-JUL-1999: 99US-0145089  
PR 22-JUL-1999: 99US-0145192

PR 23-JUL-1999: 99US-0145145  
PR 23-JUL-1999: 99US-0145218  
PR 26-JUL-1999: 99US-0145224  
PR 26-JUL-1999: 99US-0145276  
PR 27-JUL-1999: 99US-0145913  
PR 27-JUL-1999: 99US-0145918  
PR 27-JUL-1999: 99US-0145919  
PR 28-JUL-1999: 99US-0145961  
PR 02-AUG-1999: 99US-0146186  
PR 02-AUG-1999: 99US-0146388  
PR 03-AUG-1999: 99US-0146389  
PR 04-AUG-1999: 99US-0147038  
PR 04-AUG-1999: 99US-0147204  
PR 04-AUG-1999: 99US-0147302  
PR 05-AUG-1999: 99US-0147192  
PR 05-AUG-1999: 99US-0147260  
PR 06-AUG-1999: 99US-0147303  
PR 06-AUG-1999: 99US-0147416  
PR 09-AUG-1999: 99US-0147493  
PR 09-AUG-1999: 99US-0147935  
PR 10-AUG-1999: 99US-0148171  
PR 11-AUG-1999: 99US-0148319  
PR 12-AUG-1999: 99US-0148341  
PR 13-AUG-1999: 99US-0148565  
PR 13-AUG-1999: 99US-0148584  
PR 16-AUG-1999: 99US-0149368  
PR 17-AUG-1999: 99US-0149175  
PR 18-AUG-1999: 99US-0149426  
PR 20-AUG-1999: 99US-0149722  
PR 20-AUG-1999: 99US-0149723  
PR 20-AUG-1999: 99US-0149929  
PR 23-AUG-1999: 99US-0149902  
PR 24-AUG-1999: 99US-0149930  
PR 25-AUG-1999: 99US-0150566  
PR 26-AUG-1999: 99US-0150884  
PR 27-AUG-1999: 99US-0151065  
PR 27-AUG-1999: 99US-0151066  
PR 27-AUG-1999: 99US-0151080  
PR 30-AUG-1999: 99US-0151403  
PR 31-AUG-1999: 99US-0151438  
PR 01-SEP-1999: 99US-0151930  
PR 07-SEP-1999: 99US-0152363  
PR 10-SEP-1999: 99US-0153070  
PR 13-SEP-1999: 99US-0153758  
PR 15-SEP-1999: 99US-0154039  
PR 16-SEP-1999: 99US-0154039  
PR 20-SEP-1999: 99US-0154779  
PR 22-SEP-1999: 99US-0155149  
PR 23-SEP-1999: 99US-0155486  
PR 24-SEP-1999: 99US-0155659  
PR 28-SEP-1999: 99US-0156458  
PR 29-SEP-1999: 99US-0156596  
PR 04-OCT-1999: 99US-0157117  
PR 05-OCT-1999: 99US-0157753  
PR 06-OCT-1999: 99US-0157665  
PR 07-OCT-1999: 99US-0158029  
PR 08-OCT-1999: 99US-0158232  
PR 12-OCT-1999: 99US-0158369  
PR 13-OCT-1999: 99US-0159293  
PR 13-OCT-1999: 99US-0159294  
PR 13-OCT-1999: 99US-0159295  
PR 14-OCT-1999: 99US-0159429  
PR 14-OCT-1999: 99US-0159430  
PR 14-OCT-1999: 99US-0159431  
PR 14-OCT-1999: 99US-0159637  
PR 14-OCT-1999: 99US-0159638  
PR 18-OCT-1999: 99US-0159584  
PR 21-OCT-1999: 99US-0160741  
PR 21-OCT-1999: 99US-0160767  
PR 21-OCT-1999: 99US-0160768  
PR 21-OCT-1999: 99US-0160770  
PR 21-OCT-1999: 99US-0160814  
PR 21-OCT-1999: 99US-0160815

PR	05-MAR-1999;	99US-0123180.
PR	09-MAR-1999;	99US-0123548.
PR	23-MAR-1999;	99US-0125788.

FR	5	MAY-1999	9905-01-65264
FR	23	MAY-1999	9905-01-67785
FR	01	JUN-1999	9905-01-27452
FR	08	APR-1999	9905-01-28724
FR	16	APR-1999	9905-01-28714
FR	15	APR-1999	9905-01-29845
FR	23	APR-1999	9905-01-30077
FR	23	APR-1999	9905-01-30049
FR	23	APR-1999	9905-01-30510
FR	28	APR-1999	9905-01-30891
FR	30	APR-1999	9905-01-31449
FR	04	MAY-1999	9905-01-32048
FR	05	MAY-1999	9905-01-32404
FR	05	MAY-1999	9905-01-32485
FR	06	MAY-1999	9905-01-32486
FR	06	MAY-1999	9905-01-32487
FR	07	MAY-1999	9905-01-32863
FR	11	MAY-1999	9905-01-34256
FR	14	MAY-1999	9905-01-34218
FR	14	MAY-1999	9905-01-34219
FR	14	MAY-1999	9905-01-34221
FR	14	MAY-1999	9905-01-34370
FR	18	MAY-1999	9905-01-34768
FR	19	MAY-1999	9905-01-35194
FR	20	MAY-1999	9905-01-35124
FR	21	MAY-1999	9905-01-35353
FR	24	MAY-1999	9905-01-35737
FR	25	MAY-1999	9905-01-36021
FR	26	MAY-1999	9905-01-36302
FR	26	MAY-1999	9905-01-36782
FR	01	JUN-1999	9905-01-37222
FR	03	JUN-1999	9905-01-37528
FR	04	JUN-1999	9905-01-37520
FR	07	JUN-1999	9905-01-37724
FR	08	JUN-1999	9905-01-38094
FR	10	JUN-1999	9905-01-38540
FR	14	JUN-1999	9905-01-39119
FR	15	JUN-1999	9905-01-39453
FR	17	JUN-1999	9905-01-39452
FR	18	JUN-1999	9905-01-39454
FR	18	JUN-1999	9905-01-39455
FR	18	JUN-1999	9905-01-39456
FR	18	JUN-1999	9905-01-39457
FR	18	JUN-1999	9905-01-39458
FR	18	JUN-1999	9905-01-39459
FR	18	JUN-1999	9905-01-39460
FR	18	JUN-1999	9905-01-39461
FR	18	JUN-1999	9905-01-39462
FR	18	JUN-1999	9905-01-39463
FR	18	JUN-1999	9905-01-39464
FR	18	JUN-1999	9905-01-39465
FR	18	JUN-1999	9905-01-39466
FR	18	JUN-1999	9905-01-39467
FR	18	JUN-1999	9905-01-39468
FR	18	JUN-1999	9905-01-39469
FR	18	JUN-1999	9905-01-39470
FR	18	JUN-1999	9905-01-39471
FR	18	JUN-1999	9905-01-39472
FR	18	JUN-1999	9905-01-39473
FR	18	JUN-1999	9905-01-39474
FR	18	JUN-1999	9905-01-39475
FR	18	JUN-1999	9905-01-39476
FR	18	JUN-1999	9905-01-39477
FR	18	JUN-1999	9905-01-39478
FR	18	JUN-1999	9905-01-39479
FR	18	JUN-1999	9905-01-39480
FR	18	JUN-1999	9905-01-39481
FR	18	JUN-1999	9905-01-39482
FR	18	JUN-1999	9905-01-39483
FR	18	JUN-1999	9905-01-39484
FR	18	JUN-1999	9905-01-39485
FR	18	JUN-1999	9905-01-39486
FR	18	JUN-1999	9905-01-39487
FR	18	JUN-1999	9905-01-39488
FR	18	JUN-1999	9905-01-39489
FR	18	JUN-1999	9905-01-39490
FR	18	JUN-1999	9905-01-39491
FR	18	JUN-1999	9905-01-39492
FR	18	JUN-1999	9905-01-39493
FR	18	JUN-1999	9905-01-39494
FR	18	JUN-1999	9905-01-39495
FR	18	JUN-1999	9905-01-39496
FR	18	JUN-1999	9905-01-39497
FR	18	JUN-1999	9905-01-39498
FR	18	JUN-1999	9905-01-39499
FR	18	JUN-1999	9905-01-39500
FR	18	JUN-1999	9905-01-39



PR 16-JUL-1999: 99US-0144086  
 PR 19-JUL-1999: 99US-0144325  
 PR 19-JUL-1999: 99US-0144331  
 PR 19-JUL-1999: 99US-0144332  
 PR 19-JUL-1999: 99US-0144333  
 PR 19-JUL-1999: 99US-0144334  
 PR 19-JUL-1999: 99US-0144335  
 PR 20-JUL-1999: 99US-0144352  
 PR 20-JUL-1999: 99US-0144357  
 PR 20-JUL-1999: 99US-0144884  
 PR 21-JUL-1999: 99US-0144814  
 PR 21-JUL-1999: 99US-0145086  
 PR 21-JUL-1999: 99US-0145088  
 PR 22-JUL-1999: 99US-0145085  
 PR 22-JUL-1999: 99US-0145087  
 PR 22-JUL-1999: 99US-0145089  
 PR 22-JUL-1999: 99US-0145152  
 PR 23-JUL-1999: 99US-0145145  
 PR 23-JUL-1999: 99US-0145151  
 PR 23-JUL-1999: 99US-0145224  
 PR 26-JUL-1999: 99US-0145275  
 PR 27-JUL-1999: 99US-0145313  
 PR 27-JUL-1999: 99US-0145318  
 PR 27-JUL-1999: 99US-0145319  
 PR 28-JUL-1999: 99US-0145351  
 PR 02-AUG-1999: 99US-0145386  
 PR 02-AUG-1999: 99US-0145388  
 PR 03-AUG-1999: 99US-0145389  
 PR 03-AUG-1999: 99US-0147038  
 PR 04-AUG-1999: 99US-0147204  
 PR 05-AUG-1999: 99US-0147302  
 PR 05-AUG-1999: 99US-0147360  
 PR 06-AUG-1999: 99US-0147383  
 PR 06-AUG-1999: 99US-0147416  
 PR 06-AUG-1999: 99US-0147434  
 PR 09-AUG-1999: 99US-0147435  
 PR 10-AUG-1999: 99US-0148171  
 PR 11-AUG-1999: 99US-0148319  
 PR 12-AUG-1999: 99US-0148341  
 PR 13-AUG-1999: 99US-0148365  
 PR 13-AUG-1999: 99US-0148368  
 PR 16-AUG-1999: 99US-0149368  
 PR 17-AUG-1999: 99US-0149175  
 PR 18-AUG-1999: 99US-0149425  
 PR 20-AUG-1999: 99US-0149722  
 PR 20-AUG-1999: 99US-0149723  
 PR 20-AUG-1999: 99US-0149724  
 PR 23-AUG-1999: 99US-0149902  
 PR 23-AUG-1999: 99US-0149903  
 PR 25-AUG-1999: 99US-0150566  
 PR 25-AUG-1999: 99US-0150584  
 PR 27-AUG-1999: 99US-0151065  
 PR 27-AUG-1999: 99US-0151066  
 PR 27-AUG-1999: 99US-0151080  
 PR 30-AUG-1999: 99US-0151083  
 PR 31-AUG-1999: 99US-0151138  
 PR 01-SEP-1999: 99US-0151140  
 PR 07-SEP-1999: 99US-0152363  
 PR 10-SEP-1999: 99US-0153070  
 PR 13-SEP-1999: 99US-0153758  
 PR 15-SEP-1999: 99US-0154018  
 PR 16-SEP-1999: 99US-0154039  
 PR 20-SEP-1999: 99US-0154779  
 PR 22-SEP-1999: 99US-0155139  
 PR 23-SEP-1999: 99US-0155486  
 PR 24-SEP-1999: 99US-0155489  
 PR 28-SEP-1999: 99US-0156458  
 PR 29-SEP-1999: 99US-0156596  
 PR 04-OCT-1999: 99US-0157117  
 PR 05-OCT-1999: 99US-0157753  
 PR 06-OCT-1999: 99US-0157865  
 PR 07-OCT-1999: 99US-0158029

PR 08-OCT-1999: 99US-0158232  
 PR 12-OCT-1999: 99US-0158369  
 PR 13-OCT-1999: 99US-0159293  
 PR 13-OCT-1999: 99US-0159294  
 PR 13-OCT-1999: 99US-0159295  
 PR 14-OCT-1999: 99US-0159329  
 PR 14-OCT-1999: 99US-0159330  
 PR 14-OCT-1999: 99US-0159331  
 PR 14-OCT-1999: 99US-0159637  
 PR 14-OCT-1999: 99US-0159638  
 PR 14-OCT-1999: 99US-0159584  
 PR 21-OCT-1999: 99US-0160741  
 PR 21-OCT-1999: 99US-0160767  
 PR 21-OCT-1999: 99US-0160768  
 PR 21-OCT-1999: 99US-0160770  
 PR 21-OCT-1999: 99US-0160814  
 PR 21-OCT-1999: 99US-0160815  
 PR 21-OCT-1999: 99US-0160840  
 PR 21-OCT-1999: 99US-0160981  
 PR 25-OCT-1999: 99US-0161404  
 PR 25-OCT-1999: 99US-0161405  
 PR 25-OCT-1999: 99US-0161406  
 PR 26-OCT-1999: 99US-0161359  
 PR 26-OCT-1999: 99US-0161360  
 PR 26-OCT-1999: 99US-0161361  
 PR 26-OCT-1999: 99US-0161369  
 PR 28-OCT-1999: 99US-0161992  
 PR 28-OCT-1999: 99US-0161993  
 PR 29-OCT-1999: 99US-0162142

Query Match 26.0% Score 481, DE 21, Length 386,  
 Best Local Similarity 32.8%, Pred No 7,86-40,  
 Matches 129, Conservative 56, Mismatches 134, Indels 56, Gaps 12,

QY 2 MNPEDESYANCSFTQVASMAGPALENAVEITLFSRPHQA-----LNAPLQQA 55  
 Db 21 mkqgkqgdyams---qagamharsmlhleeit-eenvhnsaasppptavdgss 75  
 QY 56 GPNTPAVISTIKPMKMKKCPKELNCGTIFLDVYINLPQNDPFTFKGISSEVYGNKCEV 115  
 Db 76 gantvhlidlvkhkiskrdaagldpeltatdsldpslditlclq-llpplvsnlwne- 133  
 QY 116 PC-----YVMQVQSGFHCPLFPNLSLHVSSYVHMLTQPKLISHEALNKG 166  
 Db 134 ecdaadgnasytvaqygsstgtrlllparlidlthsaflshwtsqpssttdtsaaynrg 193  
 QY 167 ELYT---SKSTFTVNEAVTSGFRRLMILNAPSQEVVWQWVILPPTQSTISQW 222  
 Db 194 lvtlhqgsktltt---aykltlqdiadelltdaaevkkgjgallvqglstvgldq 248  
 QY 224 GSR-----FTWELAMALAHVSGILIPKELTINHSYASILEKVLIVKPGS 273  
 Db 249 gggglilqthrtgawd-----dlvrotglvaokrotgnrnyapsldtkevvdang 301  
 QY 274 FTLHLEGFELUS--VEMQENLKWVGEKELKVPAFTFELISNQGPELMKLYLK-- 328  
 Db 302 laidklvnykgsrpllvnnopddasevgrataascsvagvlycaahgcolsklstrvs 361  
 QY 329 -----FTHLVVS 335  
 Db 362 ratslhakvlnlqfhlvas 382

RESULT 13  
 AAAG46651 standard: Protein, 387 AA.  
 XX AAAG46651;  
 AC  
 XX 18-OCT-2000 (first entry)  
 DI  
 XX Arabidopsis thaliana protein fragment SEQ ID NO: 58714.

XX Protein identification: signal transduction pathway; metabolic pathway;  
KM hybridisation assay; genetic mapping; gene expression control; protein-coding  
XX termination sequence.  
XX Arabidopsis thaliana.  
XX EF1033405 A2.  
XX  
XX 06-SEP-2000.  
XX  
XX 25 FEB-2000: 2060BP-0301430  
XX  
XX 25 FEB-1999: 990S-0121825.  
XX 05-MAR-1999: 990S-0123180.  
XX 06-MAR-1999: 990S-0123548.  
XX 23 MAR-1999: 990S-0125788.  
XX 25 MAR-1999: 990S-0126264.  
XX 29-MAR-1999: 990S-0126785.  
XX 01-APR-1999: 990S-0127462.  
XX 06-APR-1999: 990S-0128234.  
XX 08-APR-1999: 990S-0128714.  
XX 16-APR-1999: 990S-0129845.  
XX 19-APR-1999: 990S-0130077.  
XX 21-APR-1999: 990S-0130449.  
XX 23-APR-1999: 990S-0130510.  
XX 24-APR-1999: 990S-0130891.  
XX 28-APR-1999: 990S-0131449.  
XX 30-APR-1999: 990S-0132049.  
XX 30-APR-1999: 990S-0132407.  
XX 04-MAY-1999: 990S-0132484.  
XX 05-MAY-1999: 990S-0132485.  
XX 06-MAY-1999: 990S-0132486.  
XX 06-MAY-1999: 990S-0132487.  
XX 07-MAY-1999: 990S-0132863.  
XX 11-MAY-1999: 990S-0134256.  
XX 14-MAY-1999: 990S-0134218.  
XX 14-MAY-1999: 990S-0134219.  
XX 14-MAY-1999: 990S-0134221.  
XX 14-MAY-1999: 990S-0134370.  
XX 18-MAY-1999: 990S-0134768.  
XX 19-MAY-1999: 990S-0134941.  
XX 20-MAY-1999: 990S-0135124.  
XX 21-MAY-1999: 990S-0135353.  
XX 24-MAY-1999: 990S-0135629.  
XX 25-MAY-1999: 990S-0136021.  
XX 27-MAY-1999: 990S-0136392.  
XX 28-MAY-1999: 990S-0136782.  
XX 01-JUN-1999: 990S-0137222.  
XX 04-JUN-1999: 990S-0137528.  
XX 04-JUN-1999: 990S-0137502.  
XX 07-JUN-1999: 990S-0137724.  
XX 08-JUN-1999: 990S-0138094.  
XX 10-JUN-1999: 990S-0138540.  
XX 10-JUN-1999: 990S-0138847.  
XX 14-JUN-1999: 990S-0139119.  
XX 16-JUN-1999: 990S-0139452.  
XX 16-JUN-1999: 990S-0139453.  
XX 17-JUN-1999: 990S-0139454.  
XX 18-JUN-1999: 990S-0139454.  
XX 18-JUN-1999: 990S-0139455.  
XX 18-JUN-1999: 990S-0139456.  
XX 18-JUN-1999: 990S-0139457.  
XX 18-JUN-1999: 990S-0139458.  
XX 18-JUN-1999: 990S-0139459.  
XX 18-JUN-1999: 990S-0139460.  
XX 18-JUN-1999: 990S-0139461.  
XX 18-JUN-1999: 990S-0139462.  
XX 18-JUN-1999: 990S-0139463.  
XX 18-JUN-1999: 990S-0139470.  
XX 18-JUN-1999: 990S-0139475.  
XX 21-JUN-1999: 990S-0139817.  
XX 22-JUN-1999: 990S-0139899.  
PR 23-JUN-1999: 990S-0140453.  
PR 23-JUN-1999: 990S-0140354.  
PR 24-JUN-1999: 990S-0140695.  
PR 28-JUN-1999: 990S-0140823.  
PR 29-JUN-1999: 990S-0140991.  
PR 30-JUN-1999: 990S-0141287.  
PR 01-JUL-1999: 990S-0141842.  
PR 01-JUL-1999: 990S-0142154.  
PR 02-JUL-1999: 990S-0142065.  
PR 06-JUL-1999: 990S-0142390.  
PR 08-JUL-1999: 990S-0142803.  
PR 09-JUL-1999: 990S-0142920.  
PR 12-JUL-1999: 990S-0142977.  
PR 13-JUL-1999: 990S-0143542.  
PR 14-JUL-1999: 990S-0143624.  
PR 15-JUL-1999: 990S-0144005.  
PR 16-JUL-1999: 990S-0144085.  
PR 16-JUL-1999: 990S-0144086.  
PR 19-JUL-1999: 990S-0144325.  
PR 19-JUL-1999: 990S-0144331.  
PR 19-JUL-1999: 990S-0144332.  
PR 19-JUL-1999: 990S-0144333.  
PR 19-JUL-1999: 990S-0144334.  
PR 19-JUL-1999: 990S-0144335.  
PR 20-JUL-1999: 990S-0144352.  
PR 20-JUL-1999: 990S-0144632.  
PR 20-JUL-1999: 990S-0144884.  
PR 21-JUL-1999: 990S-0144814.  
PR 21-JUL-1999: 990S-0145086.  
PR 21-JUL-1999: 990S-0145088.  
PR 22-JUL-1999: 990S-0145085.  
PR 22-JUL-1999: 990S-0145087.  
PR 23-JUL-1999: 990S-0145099.  
PR 23-JUL-1999: 990S-0145192.  
PR 23-JUL-1999: 990S-0145145.  
PR 23-JUL-1999: 990S-0145218.  
PR 23-JUL-1999: 990S-0145224.  
PR 26-JUL-1999: 990S-0145276.  
PR 27-JUL-1999: 990S-0145613.  
PR 27-JUL-1999: 990S-0145918.  
PR 27-JUL-1999: 990S-0145919.  
PR 28-JUL-1999: 990S-0145951.  
PR 02-AUG-1999: 990S-0146386.  
PR 02-AUG-1999: 990S-0146388.  
PR 02-AUG-1999: 990S-0146389.  
PR 03-AUG-1999: 990S-0147038.  
PR 04-AUG-1999: 990S-0147204.  
PR 04-AUG-1999: 990S-0147302.  
PR 05-AUG-1999: 990S-0147392.  
PR 05-AUG-1999: 990S-0147260.  
PR 06-AUG-1999: 990S-0147303.  
PR 06-AUG-1999: 990S-0147416.  
PR 06-AUG-1999: 990S-0147416.  
PR 09-AUG-1999: 990S-0147935.  
PR 10-AUG-1999: 990S-0148171.  
PR 11-AUG-1999: 990S-0148179.  
PR 12-AUG-1999: 990S-0148341.  
PR 13-AUG-1999: 990S-0148565.  
PR 13-AUG-1999: 990S-0148684.  
PR 16-AUG-1999: 990S-0148688.  
PR 17-AUG-1999: 990S-0149175.  
PR 18-AUG-1999: 990S-0149426.  
PR 20-AUG-1999: 990S-0149722.  
PR 20-AUG-1999: 990S-0149723.  
PR 20-AUG-1999: 990S-0149729.  
PR 23-AUG-1999: 990S-0149930.  
PR 23-AUG-1999: 990S-0149930.  
PR 25-AUG-1999: 990S-0149936.  
PR 26-AUG-1999: 990S-0150884.  
PR 27-AUG-1999: 990S-0151065.  
PR 27-AUG-1999: 990S-0151066.  
PR 27-AUG-1999: 990S-0151080.  
PR 30-AUG-1999: 990S-0151303.









GenForce version 4.5  
 Copyright (c) 1993 - 2000 GenForce, Inc.

GM nucleic - nucleic search, using SW model

Run on: August 31, 2001, 13:27:32 : Search time 67.13 seconds  
 (without alignments)  
 4024.246 Million cell updates/sec

Title: US-09-577-657A-2

Perfect score: 1427

align393430934410.1427

Sequence: 1 30100310934410.99

Scoring table: IDENTITY\_NUC

Gapop 10.0, Gapext 1.0

Searched: 324599 seqs, 9465562 residues

Total number of hits satisfying chosen parameters: 649198

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: Issued Patents, NA: \*

1: /cgn2\_6/prodata/2/ina/5A\_COMB seq \*

2: /cgn2\_6/prodata/2/ina/5A\_COMB seq \*

3: /cgn2\_6/prodata/2/ina/5A\_COMB seq \*

4: /cgn2\_6/prodata/2/ina/5A\_COMB seq \*

5: /cgn2\_6/prodata/2/ina/5A\_COMB seq \*

6: /cgn2\_6/prodata/2/ina/5A\_COMB seq \*

7: /cgn2\_6/prodata/2/ina/5A\_COMB seq \*

8: /cgn2\_6/prodata/2/ina/5A\_COMB seq \*

9: /cgn2\_6/prodata/2/ina/5A\_COMB seq \*

10: /cgn2\_6/prodata/2/ina/5A\_COMB seq \*

11: /cgn2\_6/prodata/2/ina/5A\_COMB seq \*

12: /cgn2\_6/prodata/2/ina/5A\_COMB seq \*

13: /cgn2\_6/prodata/2/ina/5A\_COMB seq \*

14: /cgn2\_6/prodata/2/ina/5A\_COMB seq \*

15: /cgn2\_6/prodata/2/ina/5A\_COMB seq \*

16: /cgn2\_6/prodata/2/ina/5A\_COMB seq \*

17: /cgn2\_6/prodata/2/ina/5A\_COMB seq \*

18: /cgn2\_6/prodata/2/ina/5A\_COMB seq \*

19: /cgn2\_6/prodata/2/ina/5A\_COMB seq \*

20: /cgn2\_6/prodata/2/ina/5A\_COMB seq \*

21: /cgn2\_6/prodata/2/ina/5A\_COMB seq \*

22: /cgn2\_6/prodata/2/ina/5A\_COMB seq \*

23: /cgn2\_6/prodata/2/ina/5A\_COMB seq \*

24: /cgn2\_6/prodata/2/ina/5A\_COMB seq \*

25: /cgn2\_6/prodata/2/ina/5A\_COMB seq \*

26: /cgn2\_6/prodata/2/ina/5A\_COMB seq \*

27: /cgn2\_6/prodata/2/ina/5A\_COMB seq \*

# SUMMARIES

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## ALIGNMENTS

28	46	3.2	1737	2	US-08-801-228-3	Sequence 3, Appl 1
29	46	3.2	1737	3	US-09-104-296-3	Sequence 3, Appl 1
30	46	3.2	1737	4	US-09-104-296-3	Sequence 3, Appl 1
31	45.5	3.2	1737	5	US-09-104-296-3	Sequence 3, Appl 1
32	45.4	3.2	1737	6	US-09-104-296-3	Sequence 3, Appl 1
33	45.2	3.2	1737	7	US-09-104-296-3	Sequence 3, Appl 1
34	45.1	3.2	1737	8	US-09-104-296-3	Sequence 3, Appl 1
35	45.0	3.2	1737	9	US-09-104-296-3	Sequence 3, Appl 1
36	45.0	3.2	1737	10	US-09-104-296-3	Sequence 3, Appl 1
37	45.0	3.2	1737	11	US-09-104-296-3	Sequence 3, Appl 1
38	44.8	3.1	1737	12	US-09-104-296-3	Sequence 3, Appl 1
39	44.6	3.1	1737	13	US-09-104-296-3	Sequence 3, Appl 1
40	44.4	3.1	1737	14	US-09-104-296-3	Sequence 3, Appl 1
41	44.2	3.1	1737	15	US-09-104-296-3	Sequence 3, Appl 1
42	44.0	3.1	1737	16	US-09-104-296-3	Sequence 3, Appl 1
43	43.6	3.1	1737	17	US-09-104-296-3	Sequence 3, Appl 1
44	43.6	3.1	1737	18	US-09-104-296-3	Sequence 3, Appl 1
45	43.6	3.1	1737	19	US-09-104-296-3	Sequence 3, Appl 1

## RESULT 1

US-09-027-137-2

Sequence 2, Application US-09027137

Patent No. 6013450

GENERAL INFORMATION:

APPLICANT: Hillman, Jennifer L.

APPLICANT: Corley, Neil C.

APPLICANT: Yoo, Henry

TITLE OF INVENTION: CAPI-RELATED PROTEIN

NUMBER OF SEQUENCES: 3

CORRESPONDENCE ADDRESS:

ATTORNEY/AGENT INFORMATION:

NAME: Billings, Lucy J.

REFERENCE/DOCKET NUMBER: 36,749

TELECOMMUNICATION INFORMATION:

TELEPHONE: 650-855-0555

TELEFAX: 650-845-4166

INFORMATION FOR SEQ. ID NO. 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 2852 base pairs

TYPE: nucleic acid

STRANDNESS: single

TOPOLOGY: linear

IMMEDIATE SOURCE:

LIBRARY: FOSNM16

CLONE: 2299466

US-09-027-137-2

Query Match: 3.7%

Best Local Similarity: 50.2%

Source: 52.6, DB: 2

Pred. No.: 0.00015

Length: 2852









Query Match	3.78;	Score 52.4;	EH 4;	Length 2504;
Best Local Similarity	49.68;	pred. No. 0.00017;		
Matches 134, Conservative	0,	Mismatches 136,	Indels 0,	Gaps 0.

## RESULT 8

5196333-3

Matches 70; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

RESULT 9  
NO. 00 157

COUNTRY: USA

US-08-157-101A-4

1012 AAAAAAAA 1019





```

1 RESULT 15
2 US-09-064-950-1
3 Sequence 1, Application US/09064950C
4 Patent No. 6225085
5
6 GENERAL INFORMATION:
7
8 APPLICANT: Holtzman, Douglas A.
9
10 TITLE OF INVENTION: NINE1 LRSQ PROTEIN AND NUCLEIC ACID MOLECULES AND USES
11
12 TITLE OF INVENTION: THEREFOR
13
14 FILE REFERENCE: ME1-019
15
16 CURRENT APPLICATION NUMBER: US/09/063,950C
17
18 CURRENT FILING DATE: 1998-04-21
19
20 NUMBER OF SEQ ID NOS: 9
21
22 SOFTWARE: PatentIn Ver. 2.0
23
24 SEQ ID NO. 1
25
26 LENGTH: 2852
27
28 TYPE: DNA
29
30 ORGANISM: Homo sapiens
31
32 FEATURE:

```

Qy	1410	tcctgcgcgcgc	1421
Dy	2841	aaagcgcgcgcgc	2852

Search completed: August 31, 2001, 14:22:18  
Job time: 3286 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 31, 2001, 13:15:51: Search time 1159.97 seconds  
(without alignments)  
11628.940 Million cell updates/sec

Title: US-09-577-657A-2

Perfect score: 1427

Sequence: 1 tgaatacagctgctgtgtgag.....atctgcgagcgcgaattc 1427

Scoring table:

IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Searched: 10228115 seqs, 4726426750 residues

Total number of hits satisfying chosen parameters: 20456230

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

EST: \*

1: qb\_est1: \*  
2: qb\_est12: \*  
3: qb\_est13: \*  
4: qb\_est14: \*  
5: qb\_est15: \*  
6: qb\_est16: \*  
7: qb\_est17: \*  
8: qb\_est18: \*  
9: qb\_est19: \*  
10: qb\_est10: \*  
11: qb\_est11: \*  
12: qb\_est12: \*  
13: qb\_est13: \*  
14: qb\_est14: \*  
15: qb\_est15: \*  
16: qb\_est16: \*  
17: qb\_est17: \*  
18: qb\_est18: \*  
19: qb\_est19: \*  
20: qb\_est20: \*  
21: qb\_est21: \*  
22: qb\_est22: \*  
23: qb\_est23: \*  
24: qb\_est24: \*  
25: qb\_est25: \*  
26: qb\_est26: \*  
27: qb\_est27: \*  
28: qb\_est28: \*  
29: qb\_est29: \*  
30: qb\_est30: \*  
31: qb\_est31: \*  
32: qb\_est32: \*  
33: qb\_est33: \*  
34: qb\_est34: \*  
35: qb\_est35: \*  
36: qb\_est36: \*  
37: qb\_est37: \*  
38: qb\_est38: \*  
39: qb\_est39: \*  
40: qb\_est40: \*  
41: qb\_est41: \*  
42: qb\_est42: \*  
43: qb\_est43: \*

44: em\_esthum10: \*  
45: em\_esthum11: \*  
46: em\_esthum12: \*  
47: em\_esthum13: \*  
48: em\_esthum14: \*  
49: em\_esthum15: \*  
50: em\_esthum16: \*  
51: em\_esthum17: \*  
52: em\_esthum18: \*  
53: em\_esthum19: \*  
54: em\_esthum20: \*  
55: em\_esthum21: \*  
56: em\_esthum22: \*  
57: em\_esthum23: \*  
58: em\_esthum24: \*  
59: em\_esthum25: \*  
60: em\_esthum26: \*  
61: em\_esthum27: \*  
62: em\_esthum28: \*  
63: em\_esthum29: \*  
64: em\_esthum30: \*  
65: em\_esthum31: \*  
66: em\_esthum32: \*  
67: em\_esthum33: \*  
68: em\_esthum34: \*  
69: em\_esthum35: \*  
70: em\_esthum36: \*  
71: em\_esthum37: \*  
72: em\_esthum38: \*  
73: em\_esthum39: \*  
74: em\_esthum40: \*  
75: em\_esthum41: \*  
76: em\_esthum42: \*  
77: em\_esthum43: \*  
78: em\_esthum44: \*  
79: em\_esthum45: \*  
80: em\_esthum46: \*  
81: em\_esthum47: \*  
82: em\_esthum48: \*  
83: em\_esthum49: \*  
84: em\_esthum50: \*  
85: em\_esthum51: \*  
86: em\_esthum52: \*  
87: em\_esthum53: \*  
88: em\_esthum54: \*  
89: em\_esthum55: \*  
90: em\_esthum56: \*  
91: em\_esthum57: \*  
92: em\_esthum58: \*  
93: em\_esthum59: \*  
94: em\_esthum60: \*  
95: em\_esthum61: \*  
96: em\_esthum62: \*  
97: em\_esthum63: \*  
98: em\_esthum64: \*  
99: em\_esthum65: \*  
100: em\_esthum66: \*  
101: em\_esthum67: \*  
102: em\_esthum68: \*  
103: em\_esthum69: \*  
104: em\_esthum70: \*  
105: em\_esthum71: \*  
106: em\_esthum72: \*  
107: em\_esthum73: \*  
108: em\_esthum74: \*  
109: em\_esthum75: \*  
110: em\_esthum76: \*  
111: em\_esthum77: \*  
112: em\_esthum78: \*  
113: em\_esthum79: \*  
114: em\_esthum80: \*  
115: em\_esthum81: \*  
116: em\_esthum82: \*

```

117: qb_est148:*
118: qb_est149:*
119: qb_est150:*
120: qb_est151:*
121: qb_est152:*
122: qb_est153:*
123: qb_est154:*
124: qb_est155:*
125: qb_est156:*
126: qb_est157:*
127: qb_est158:*
128: qb_est159:*
129: qb_est160:*
130: qb_est161:*
131: qb_est162:*
132: qb_est163:*
133: qb_est164:*
134: qb_est165:*
135: qb_est166:*
136: qb_est167:*
137: qb_est168:*
138: qb_est169:*
139: qb_est170:*
140: qb_est171:*
141: qb_est172:*
142: qb_est173:*
143: qb_est174:*
144: qb_est175:*
145: qb_est176:*
146: qb_est177:*
147: qb_est178:*
148: qb_est179:*
149: qb_est180:*
150: qb_est181:*
151: qb_est182:*
152: qb_est183:*
153: qb_est184:*
154: qb_est185:*
155: qb_est186:*
156: qb_est187:*
157: qb_est188:*
158: qb_est189:*
159: qb_est190:*
160: qb_est191:*
161: qb_est192:*
162: qb_est193:*
163: qb_est194:*
164: qb_est195:*
165: qb_est196:*
166: qb_est197:*
167: qb_est198:*
168: qb_est199:*
169: qb_est200:*
170: qb_est201:*
171: qb_est202:*
172: qb_est203:*
173: qb_est204:*
174: qb_est205:*
175: qb_est206:*
176: qb_est207:*
177: qb_est208:*
178: qb_est209:*
179: qb_est210:*
180: qb_est211:*
181: qb_est212:*
182: qb_est213:*
183: qb_est214:*
184: qb_est215:*
185: qb_est216:*
186: qb_est217:*
187: qb_est218:*
188: qb_est219:*
189: qb_est220:*

```

---

```

190: qb_est221:*
191: qb_est222:*
192: qb_est223:*
193: qb_est224:*
194: qb_est225:*
195: qb_est226:*
196: qb_est227:*
197: qb_est228:*
198: qb_est229:*
199: qb_est230:*
200: qb_est231:*
201: qb_est232:*
202: qb_est233:*
203: qb_est234:*
204: qb_est235:*
205: qb_est236:*
206: qb_est237:*
207: qb_est238:*
208: qb_est239:*
209: qb_est240:*
210: qb_est241:*
211: qb_est242:*
212: qb_est243:*
213: qb_est244:*
214: qb_est245:*
215: qb_est246:*
216: qb_est247:*
217: qb_est248:*
218: qb_est249:*
219: qb_est250:*
220: qb_est251:*
221: qb_est252:*
222: qb_est253:*
223: qb_est254:*
224: qb_est255:*
225: qb_est256:*
226: qb_est257:*
227: qb_est258:*
228: qb_est259:*
229: qb_est260:*
230: qb_est261:*
231: qb_est262:*
232: qb_est263:*
233: qb_est264:*
234: qb_est265:*
235: qb_est266:*
236: qb_est267:*
237: qb_est268:*
238: qb_est269:*
239: qb_est270:*
240: qb_est271:*
241: qb_est272:*
242: qb_est273:*
243: qb_est274:*
244: qb_est275:*
245: qb_est276:*
246: qb_est277:*
247: qb_est278:*
248: qb_est279:*
249: qb_est280:*
250: qb_est281:*
251: qb_est282:*
252: qb_est283:*
253: qb_est284:*
254: qb_est285:*
255: qb_est286:*
256: qb_est287:*
257: qb_est288:*
258: qb_est289:*

```

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

























Sat Sep 11:45:40 2001

us-09-577-657a-2.rst

Page 13

Job time: 1964 sec

---

...

GenCore version 4.5  
Copyright (c) 1994 2000 Compugen Ltd

QM nucleotide - nucleotide search, using SW model

Run on: August 31, 2001, 13:28:02 Search time 115.38 seconds

(without alignments)  
7765.779 Million cell updates/sec

Title: US-09-577-657A-2

Sequence: 1 tgcataatcgtgtgtgag...atctctggggcggcgaattc 1427

Scoring table: IDENTITY\_NUC

Gapop 10.0, Gapext 1.0

Searched: 730161 seqs, 415500000 residues

Total number of hits satisfying chosen parameters: 1460402

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: N\_Geneseq.0601.\*

- 1: /SITS1/gcgdata/geneseq/geneseq/NA1941.DAT.\*
- 2: /SITS1/gcgdata/geneseq/geneseq/NA1941.DAT.\*
- 3: /SITS1/gcgdata/geneseq/geneseq/NA1982.DAT.\*
- 4: /SITS1/gcgdata/geneseq/geneseq/NA1983.DAT.\*
- 5: /SITS1/gcgdata/geneseq/geneseq/NA1984.DAT.\*
- 6: /SITS1/gcgdata/geneseq/geneseq/NA1985.DAT.\*
- 7: /SITS1/gcgdata/geneseq/geneseq/NA1986.DAT.\*
- 8: /SITS1/gcgdata/geneseq/geneseq/NA1987.DAT.\*
- 9: /SITS1/gcgdata/geneseq/geneseq/NA1988.DAT.\*
- 10: /SITS1/gcgdata/geneseq/geneseq/NA1989.DAT.\*
- 11: /SITS1/gcgdata/geneseq/geneseq/NA1990.DAT.\*
- 12: /SITS1/gcgdata/geneseq/geneseq/NA1991.DAT.\*
- 13: /SITS1/gcgdata/geneseq/geneseq/NA1992.DAT.\*
- 14: /SITS1/gcgdata/geneseq/geneseq/NA1993.DAT.\*
- 15: /SITS1/gcgdata/geneseq/geneseq/NA1994.DAT.\*
- 16: /SITS1/gcgdata/geneseq/geneseq/NA1995.DAT.\*
- 17: /SITS1/gcgdata/geneseq/geneseq/NA1996.DAT.\*
- 18: /SITS1/gcgdata/geneseq/geneseq/NA1997.DAT.\*
- 19: /SITS1/gcgdata/geneseq/geneseq/NA1998.DAT.\*
- 20: /SITS1/gcgdata/geneseq/geneseq/NA1999.DAT.\*
- 21: /SITS1/gcgdata/geneseq/geneseq/NA2000.DAT.\*
- 22: /SITS1/gcgdata/geneseq/geneseq/NA2001.DAT.\*

Pred No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	FR	ID	Description
1	1427	100.0	1427	22	AAFC82742	C. sinensis N-meth
2	1427	100.0	1427	22	AAFC82743	C. sinensis N-meth
3	170.4	11.9	1239	21	AAFC45306	Arabidopsis thalia
4	170.4	11.9	1407	21	AAFC45421	Arabidopsis thalia
5	152.8	10.7	1445	21	AAFC45444	Arabidopsis thalia
6	130	9.1	936	22	AAFC8254	oligonucleotide D1
7	130	9.1	936	22	AAFC8254	oligonucleotide D1
8	130	9.1	936	22	AAFC8257	oligonucleotide D1
9	130	9.1	936	22	AAFC8259	oligonucleotide D2
10	130	9.1	936	22	AAFC8262	oligonucleotide D2
11	130	9.1	938	22	AAFC8255	oligonucleotide D1

C	12	129	9.0	936	22	AAFC8252	oligonucleotide D1
C	13	129	9.0	936	22	AAFC8254	oligonucleotide D1
C	14	129	9.0	936	22	AAFC8257	oligonucleotide D1
C	15	129	9.0	936	22	AAFC8259	oligonucleotide D2
C	16	129	9.0	936	22	AAFC8262	oligonucleotide D2
C	17	129	9.0	938	22	AAFC8255	oligonucleotide D1
C	18	72.8	5.1	1211	21	AAFC3567	Arabidopsis thalia
C	19	72.2	5.0	1231	21	AAFC4528	Arabidopsis thalia
C	20	71.2	5.0	1206	21	AAFC4873	Arabidopsis thalia
C	21	71.2	5.0	1259	21	AAFC4877	Arabidopsis thalia
C	22	69	4.8	1261	21	AAFC3651	Arabidopsis thalia
C	23	65.4	4.7	1273	22	AAFC3648	Arabidopsis thalia
C	24	63.4	4.7	1244	22	AAFC6238	oligonucleotide D1
C	25	62.8	4.7	1322	19	AAV43616	Human secreted pro
C	26	62.4	4.7	244	22	AAFC8238	oligonucleotide D1
C	27	62.4	4.7	2481	18	AAFC8748	pat neurotrophin Y
C	28	62.4	4.7	2481	20	AAFC8195	pat neurotrophin Y
C	29	62.4	4.7	2604	18	AAFC7595	pat neurotrophin Y
C	30	62.4	4.7	2604	20	AAFC8196	pat neurotrophin Y
C	31	62.2	4.7	454	21	AAFC3713	Arabidopsis thalia
C	32	62.2	4.7	1350	20	AAFC3202	Arabidopsis thalia
C	33	52.2	3.7	2057	21	AAFC7876	Fusarium venenatum
C	34	51.2	3.6	282	21	AAFC0972	Fusarium venenatum
C	35	50.4	3.5	585	21	AAFC08584	Human secreted pro
C	36	50.4	3.5	2636	21	AAFC69534	Human secreted pro
C	37	50.4	3.5	2636	21	AAFC69577	Human secreted pro
C	38	50.4	3.5	2636	21	AAFC69577	Human secreted pro
C	39	50.4	3.5	2636	22	AAFC24180	Human secreted pro
C	40	50.4	3.5	2880	21	AAV43438	Human secreted pro
C	41	50	3.5	1711	19	AAV43617	Human secreted pro
C	42	50	3.5	1554	21	AAFC3441	Human secreted pro
C	43	49.8	3.5	1443	22	AAFC3232	Human secreted pro
C	44	49.4	3.5	910	21	AAFC5708	Human secreted pro
C	45	49.4	3.5	2564	22	AAFC5127	Human TANC2 273 cd

#### ALIGNMENTS

RESULT	1
AAFC82742	standard: DNA: 1427 bp
AC	AAFC82742:
XX	1b-MAY-2001 (first entry)
XX	C. sinensis N-methyl transferase cDNA.
XX	Caffeine biosynthesis, N-methyl transferase, tea, N-methyl, plant; ds.
XX	Camellia sinensis.
XX	EP1055727-A2.
XX	29-MAY-2000.
XX	26-MAY-2000: 2900FP 0204522.
XX	26-MAY-1999: 99JP-0146358.
XX	(MITA) MITSUI CHEM INC.
XX	Mizuno M, Ashihara H, Mizuno K, Fujimura T;
XX	Witt, 2001 07-01-08.
XX	P-PSDB: AAB45799.
XX	DNA encoding N-methyl transferase with enzyme activities of 7 methyl
XX	transferase, N-methyl transferase, Arabidopsis N-methyl transferase and
XX	Arabidopsis N-methyl transferase, useful for caffeine synthesis in
XX	plants or microorganisms





FD 06-SEP-2000.  
XX  
XX 25-FEB-2000: 2000P: 03014 49  
PR 25-FEB-1999: 990S-0121825.  
PR 05-MAR-1999: 990S-0123180.  
PR 09-MAR-1999: 990S-0123548.  
PR 23-MAR-1999: 990S-0125788.  
PR 25-MAR-1999: 990S-0126264.  
PR 29-MAR-1999: 990S-0126785.  
PR 01-APR-1999: 990S-0127462.  
PR 06-APR-1999: 990S-0128234.  
PR 08-APR-1999: 990S-0128714.  
PR 16-APR-1999: 990S-0129845.  
PR 19-APR-1999: 990S-0130077.  
PR 21-APR-1999: 990S-0130449.  
PR 23-APR-1999: 990S-0130510.  
PR 23-APR-1999: 990S-0130891.  
PR 28-APR-1999: 990S-0131449.  
PR 30-APR-1999: 990S-0132048.  
PR 04-APR-1999: 990S-0132407.  
PR 04-MAY-1999: 990S-0132484.  
PR 05-MAY-1999: 990S-0132485.  
PR 06-MAY-1999: 990S-0132486.  
PR 06-MAY-1999: 990S-0132487.  
PR 07-MAY-1999: 990S-0132863.  
PR 11-MAY-1999: 990S-0134256.  
PR 14-MAY-1999: 990S-0134218.  
PR 14-MAY-1999: 990S-0134219.  
PR 14-MAY-1999: 990S-0134221.  
PR 14-MAY-1999: 990S-0134370.  
PR 18-MAY-1999: 990S-0134768.  
PR 19-MAY-1999: 990S-0134941.  
PR 20-MAY-1999: 990S-0135124.  
PR 21-MAY-1999: 990S-0135353.  
PR 24-MAY-1999: 990S-0135629.  
PR 25-MAY-1999: 990S-0136021.  
PR 27-MAY-1999: 990S-0136392.  
PR 28-MAY-1999: 990S-0136782.  
PR 01-JUN-1999: 990S-0137222.  
PR 03-JUN-1999: 990S-0137228.  
PR 04-JUN-1999: 990S-0137502.  
PR 07-JUN-1999: 990S-0137724.  
PR 08-JUN-1999: 990S-0138094.  
PR 10-JUN-1999: 990S-0138540.  
PR 10-JUN-1999: 990S-0138847.  
PR 14-JUN-1999: 990S-0139119.  
PR 16-JUN-1999: 990S-0139452.  
PR 16-JUN-1999: 990S-0139453.  
PR 17-JUN-1999: 990S-0139492.  
PR 18-JUN-1999: 990S-0139454.  
PR 18-JUN-1999: 990S-0139455.  
PR 18-JUN-1999: 990S-0139456.  
PR 18-JUN-1999: 990S-0139457.  
PR 18-JUN-1999: 990S-0139458.  
PR 18-JUN-1999: 990S-0139459.  
PR 18-JUN-1999: 990S-0139460.  
PR 18-JUN-1999: 990S-0139461.  
PR 18-JUN-1999: 990S-0139462.  
PR 18-JUN-1999: 990S-0139463.  
PR 18-JUN-1999: 990S-0139750.  
PR 18-JUN-1999: 990S-0139763.  
PR 21-JUN-1999: 990S-0139817.  
PR 22-JUN-1999: 990S-0139899.  
PR 23-JUN-1999: 990S-0140353.  
PR 23-JUN-1999: 990S-0140354.  
PR 24-JUN-1999: 990S-0140695.  
PR 28-JUN-1999: 990S-0140823.  
PR 29-JUN-1999: 990S-0140991.  
PR 30-JUN-1999: 990S-0141287.  
PR 01-JUL-1999: 990S-0141842.  
PR 01-JUL-1999: 990S-0142154.  
PR 02-JUL-1999: 990S-0142055.  
PR 06-JUL-1999: 990S-0142360.  
PR 08-JUL-1999: 990S-0142803.  
PR 09-JUL-1999: 990S-0142920.  
PR 12-JUL-1999: 990S-0142977.  
PR 13-JUL-1999: 990S-0143542.  
PR 14-JUL-1999: 990S-0143624.  
PR 15-JUL-1999: 990S-0144005.  
PR 16-JUL-1999: 990S-0144085.  
PR 16-JUL-1999: 990S-0144086.  
PR 16-JUL-1999: 990S-0144125.  
PR 16-JUL-1999: 990S-0144125.  
PR 19-JUL-1999: 990S-0144342.  
PR 19-JUL-1999: 990S-0144343.  
PR 19-JUL-1999: 990S-0144344.  
PR 19-JUL-1999: 990S-0144344.  
PR 20-JUL-1999: 990S-0144345.  
PR 20-JUL-1999: 990S-0144352.  
PR 20-JUL-1999: 990S-0144432.  
PR 20-JUL-1999: 990S-0144484.  
PR 21-JUL-1999: 990S-0144414.  
PR 21-JUL-1999: 990S-0144508.  
PR 21-JUL-1999: 990S-0145086.  
PR 22-JUL-1999: 990S-0145085.  
PR 22-JUL-1999: 990S-0145087.  
PR 22-JUL-1999: 990S-0145089.  
PR 22-JUL-1999: 990S-0145192.  
PR 23-JUL-1999: 990S-0145145.  
PR 23-JUL-1999: 990S-0145218.  
PR 23-JUL-1999: 990S-0145224.  
PR 23-JUL-1999: 990S-0145276.  
PR 23-JUL-1999: 990S-0145276.  
PR 27-JUL-1999: 990S-0145918.  
PR 27-JUL-1999: 990S-0145919.  
PR 28-JUL-1999: 990S-0145951.  
PR 02-AUG-1999: 990S-0146386.  
PR 02-AUG-1999: 990S-0146388.  
PR 02-AUG-1999: 990S-0146389.  
PR 03-AUG-1999: 990S-0147048.  
PR 04-AUG-1999: 990S-0147204.  
PR 04-AUG-1999: 990S-0147402.  
PR 05-AUG-1999: 990S-0147132.  
PR 05-AUG-1999: 990S-0147260.  
PR 06-AUG-1999: 990S-0147303.  
PR 06-AUG-1999: 990S-0147416.  
PR 06-AUG-1999: 990S-0147416.  
PR 06-AUG-1999: 990S-0147493.  
PR 10-AUG-1999: 990S-0147935.  
PR 11-AUG-1999: 990S-0148171.  
PR 12-AUG-1999: 990S-0148319.  
PR 12-AUG-1999: 990S-0148341.  
PR 13-AUG-1999: 990S-0148565.  
PR 13-AUG-1999: 990S-0148684.  
PR 16-AUG-1999: 990S-0148684.  
PR 17-AUG-1999: 990S-0149175.  
PR 18-AUG-1999: 990S-0149426.  
PR 20-AUG-1999: 990S-0149722.  
PR 20-AUG-1999: 990S-0149723.  
PR 20-AUG-1999: 990S-0149723.  
PR 20-AUG-1999: 990S-0149723.  
PR 20-AUG-1999: 990S-0149902.  
PR 23-AUG-1999: 990S-0149930.  
PR 25-AUG-1999: 990S-0150666.  
PR 26-AUG-1999: 990S-0150884.  
PR 27-AUG-1999: 990S-0151065.  
PR 27-AUG-1999: 990S-0151066.  
PR 27-AUG-1999: 990S-0151080.  
PR 30-AUG-1999: 990S-0151303.  
PR 31-AUG-1999: 990S-0151348.  
PR 01-SEP-1999: 990S-0151930.  
PR 07-SEP-1999: 990S-0152363.  
PR 10-SEP-1999: 990S-0153070.  
PR 13-SEP-1999: 990S-0153758.  
PR 15-SEP-1999: 990S-0154018.  
PR 16-SEP-1999: 990S-0154039.  
PR 20-SEP-1999: 990S-0154779.  
PR 22-SEP-1999: 990S-0155139.





PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 24-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 04-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 06-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 07-MAY-1999; 99US-0132487.  
PR 11-MAY-1999; 99US-0132863.  
PR 14-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137322.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140153.  
PR 23-JUN-1999; 99US-0140354.  
PR 23-JUN-1999; 99US-0140355.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140891.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142366.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142820.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144353.  
PR 20-JUL-1999; 99US-0144354.  
PR 21-JUL-1999; 99US-0144484.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145092.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0146113.  
PR 27-JUL-1999; 99US-0146388.  
PR 27-JUL-1999; 99US-0146389.  
PR 27-JUL-1999; 99US-0146390.  
PR 28-JUL-1999; 99US-0146391.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147202.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147322.  
PR 05-AUG-1999; 99US-0147360.  
PR 06-AUG-1999; 99US-0147403.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147494.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148365.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149724.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151067.  
PR 30-AUG-1999; 99US-0151080.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156506.  
PR 04-OCT-1999; 99US-0157173.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0158293.  
PR 13-OCT-1999; 99US-0159294.

```

PR 13-OCT-1999: 990S-0159295
PR 14-OCT-1999: 990S-0159429
PR 14-OCT-1999: 990S-0159430
PR 14-OCT-1999: 990S-0159331
PR 14-OCT-1999: 990S-0159637
PR 14-OCT-1999: 990S-0159638
PR 18-OCT-1999: 990S-0159584
PR 21-OCT-1999: 990S-0160741
PR 21-OCT-1999: 990S-0160767
PR 21-OCT-1999: 990S-0160768
PR 21-OCT-1999: 990S-0160770
PR 21-OCT-1999: 990S-0160814
PR 21-OCT-1999: 990S-0160815
PR 22-OCT-1999: 990S-0160980
PR 22-OCT-1999: 990S-0160982
PR 22-OCT-1999: 990S-0160989
PR 25-OCT-1999: 990S-0161404
PR 25-OCT-1999: 990S-0161405
PR 25-OCT-1999: 990S-0161406
PR 26-OCT-1999: 990S-0161359
PR 26-OCT-1999: 990S-0161360
PR 26-OCT-1999: 990S-0161361
PR 28-OCT-1999: 990S-0161420
PR 28-OCT-1999: 990S-0161992
PR 28-OCT-1999: 990S-0161994
PR 28-OCT-1999: 990S-0162142
PR 29-OCT-1999: 990S-0162142

```

```

Query Match: 11.38, Score 170.4, DB 21, Length 1407,
Best Local Similarity 49.7%, Prod No 2 26-29;
Matches 534; Conservative 0; Mismatches 516; Indels 24; Gaps 3;

```

```

QY 91 aagtggttgcacgaaggggaggaaggaaggtatgacacaaactcttcttca 150
DB 181 aaggtctatgtatgtatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 240
QY 151 cgaacaggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 210
DB 241 aagaagaaggttatacatgacgaacacacgtcttctgtgaagaacacagaagaatgatba 300
QY 211 cgaaggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 270
DB 301 tgaaggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 360
QY 271 aacatctgcagatgattctacacacacacacacacacacacacacacacac 330
DB 361 acctctttttgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 420
QY 431 atgacacacacacacacacacacacacacacacacacacacacacacacac 390
DB 421 acaaaacacacacacacacacacacacacacacacacacacacacacacac 480
QY 491 cgcctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 447
DB 481 cgaacttttaagttgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 540
QY 448 gtatgtatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatga 507
DB 541 gtctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 600
QY 508 attatgtatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 567
DB 601 attatgtatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 660
QY 548 gtcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 627
DB 661 attatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 702
QY 628 tttgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 687
DB 703 gtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 762
QY 688 gatcccaagaggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 747
DB 747 gatcccaagaggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 806

```

```

DB 743 gttctgaagaatgtctctctctctctctctctctctctctctctctctctctct 822
QY 748 ct---gattctcagacatgacagatgcttacttgggaactatagatgacatg 804
DB 823 ttaaggtatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 882
QY 805 ctgaatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 864
DB 883 gtaacatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 942
QY 865 gttatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 924
DB 943 ttatgtatcgaacatgacacacacacacacacacacacacacacacacacacacacacacacac 1002
QY 925 ttatgtatcgaacatgacacacacacacacacacacacacacacacacacacacacacacacac 984
DB 1003 ttaaggtatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1062
QY 985 ttgaaggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1044
DB 1063 ttaaggtatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1122
QY 1045 agttgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1104
DB 1123 atttggagaagaattatatacactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1182
QY 1105 atttggagaagaattatatacactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1168
DB 1183 atttggagaagaattatatacactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1236

```

## RESULT 5

AAC39544

ID AAC39544 standard. DNA. 1345 BP.

XX AAC39544:

DT 17-OCT-2000 (first entry)

XX Arabidopsis thaliana DNA fragment SEQ ID NO: 25021.

XX Hybridisation assay; genomic mapping; gene expression control;

KW protein identification; signal transduction pathway;

XX metabolic pathway; promoter; termination sequence; ss.

OS Arabidopsis thaliana.

FN EP1033405-A2.

XX PD 06-SEP-2000.

XX 25-FEB-2000, ZOOUP-0301439.

XX 25-FEB-1999: 990S-0121425.

XX 05-MAR-1999: 990S-0123180.

XX 09-MAR-1999: 990S-0123548.

XX 23-MAR-1999: 990S-0125788.

XX 23-MAR-1999: 990S-0127764.

XX 29-MAR-1999: 990S-0126785.

XX 01-APR-1999: 990S-0127462.

XX 06-APR-1999: 990S-0128234.

XX 08-APR-1999: 990S-0128714.

XX 16-APR-1999: 990S-0129845.

XX 18-APR-1999: 990S-0130077.

XX 21-APR-1999: 990S-0130449.

XX 23-APR-1999: 990S-0130510.

XX 23-APR-1999: 990S-0130801.

XX 28-APR-1999: 990S-0131449.

XX 30-APR-1999: 990S-0132048.

XX 30-APR-1999: 990S-0132407.

XX 04-MAY-1999: 990S-0132484.

XX 05-MAY-1999: 990S-0132485.

XX 06-MAY-1999: 990S-0132486.

```

PK 06-MAY-1999: 990S-0132487.
PK 07-MAY-1999: 990S-0132863.
PK 11-MAY-1999: 990S-0134256.
PK 14-MAY-1999: 990S-0134218.
PK 14-MAY-1999: 990S-0134219.
PK 14-MAY-1999: 990S-0134221.
PK 14-MAY-1999: 990S-0134370.
PK 14-MAY-1999: 990S-0134768.
PK 14-MAY-1999: 990S-0134941.
PK 20-MAY-1999: 990S-0135124.
PK 21-MAY-1999: 990S-0135353.
PK 24-MAY-1999: 990S-0135629.
PK 25-MAY-1999: 990S-0136021.
PK 27-MAY-1999: 990S-0136392.
PK 28-MAY-1999: 990S-0136782.
PK 01-JUN-1999: 990S-0137222.
PK 03-JUN-1999: 990S-0137528.
PK 04-JUN-1999: 990S-0137502.
PK 07-JUN-1999: 990S-0137724.
PK 08-JUN-1999: 990S-0138094.
PK 10-JUN-1999: 990S-0138540.
PK 10-JUN-1999: 990S-0138847.
PK 14-JUN-1999: 990S-0139119.
PK 16-JUN-1999: 990S-0139452.
PK 16-JUN-1999: 990S-0139453.
PK 17-JUN-1999: 990S-0139492.
PK 18-JUN-1999: 990S-0139454.
PK 18-JUN-1999: 990S-0139455.
PK 18-JUN-1999: 990S-0139456.
PK 18-JUN-1999: 990S-0139457.
PK 18-JUN-1999: 990S-0139458.
PK 18-JUN-1999: 990S-0139459.
PK 18-JUN-1999: 990S-0139460.
PK 18-JUN-1999: 990S-0139461.
PK 18-JUN-1999: 990S-0139462.
PK 18-JUN-1999: 990S-0139463.
PK 18-JUN-1999: 990S-0139750.
PK 18-JUN-1999: 990S-0139763.
PK 21-JUN-1999: 990S-0139817.
PK 22-JUN-1999: 990S-0139899.
PK 23-JUN-1999: 990S-0140353.
PK 23-JUN-1999: 990S-0140354.
PK 24-JUN-1999: 990S-0140354.
PK 24-JUN-1999: 990S-0140695.
PK 28-JUN-1999: 990S-0140823.
PK 29-JUN-1999: 990S-0140991.
PK 30-JUN-1999: 990S-0141287.
PK 01-JUL-1999: 990S-0141842.
PK 01-JUL-1999: 990S-0142154.
PK 02-JUL-1999: 990S-0142065.
PK 06-JUL-1999: 990S-0142390.
PK 08-JUL-1999: 990S-0142803.
PK 09-JUL-1999: 990S-0142920.
PK 12-JUL-1999: 990S-0142977.
PK 13-JUL-1999: 990S-0143542.
PK 14-JUL-1999: 990S-0143624.
PK 15-JUL-1999: 990S-0144005.
PK 16-JUL-1999: 990S-0144086.
PK 16-JUL-1999: 990S-0144086.
PK 19-JUL-1999: 990S-0144325.
PK 19-JUL-1999: 990S-0144331.
PK 19-JUL-1999: 990S-0144331.
PK 19-JUL-1999: 990S-0144332.
PK 19-JUL-1999: 990S-0144333.
PK 19-JUL-1999: 990S-0144334.
PK 19-JUL-1999: 990S-0144335.
PK 20-JUL-1999: 990S-0144352.
PK 20-JUL-1999: 990S-0144632.
PK 20-JUL-1999: 990S-0144684.
PK 21-JUL-1999: 990S-0144814.
PK 21-JUL-1999: 990S-0145086.
PK 21-JUL-1999: 990S-0145088.
PK 22-JUL-1999: 990S-0145085.
PK 22-JUL-1999: 990S-0145087.
PK 22-JUL-1999: 990S-0145089.

PK 22-JUL-1999: 990S-0145192.
PK 23-JUL-1999: 990S-0145145.
PK 23-JUL-1999: 990S-0145218.
PK 23-JUL-1999: 990S-0145324.
PK 26-JUL-1999: 990S-0145376.
PK 27-JUL-1999: 990S-0145913.
PK 27-JUL-1999: 990S-0145918.
PK 27-JUL-1999: 990S-0145919.
PK 28-JUL-1999: 990S-0145951.
PK 02-AUG-1999: 990S-0146386.
PK 02-AUG-1999: 990S-0146388.
PK 03-AUG-1999: 990S-0146389.
PK 03-AUG-1999: 990S-0147038.
PK 04-AUG-1999: 990S-0147204.
PK 04-AUG-1999: 990S-0147302.
PK 05-AUG-1999: 990S-0147192.
PK 06-AUG-1999: 990S-0147260.
PK 06-AUG-1999: 990S-0147303.
PK 06-AUG-1999: 990S-0147416.
PK 09-AUG-1999: 990S-0147493.
PK 09-AUG-1999: 990S-0147935.
PK 10-AUG-1999: 990S-0148171.
PK 11-AUG-1999: 990S-0148319.
PK 12-AUG-1999: 990S-0148341.
PK 13-AUG-1999: 990S-0148565.
PK 13-AUG-1999: 990S-0148684.
PK 16-AUG-1999: 990S-0149368.
PK 17-AUG-1999: 990S-0149175.
PK 18-AUG-1999: 990S-0149426.
PK 20-AUG-1999: 990S-0149722.
PK 20-AUG-1999: 990S-0149723.
PK 20-AUG-1999: 990S-0149929.
PK 23-AUG-1999: 990S-0149930.
PK 23-AUG-1999: 990S-0150506.
PK 26-AUG-1999: 990S-0150884.
PK 27-AUG-1999: 990S-0151065.
PK 27-AUG-1999: 990S-0151066.
PK 27-AUG-1999: 990S-0151080.
PK 30-AUG-1999: 990S-0151303.
PK 31-AUG-1999: 990S-0151303.
PK 01-SEP-1999: 990S-0151930.
PK 07-SEP-1999: 990S-0152363.
PK 10-SEP-1999: 990S-0153070.
PK 13-SEP-1999: 990S-0153758.
PK 15-SEP-1999: 990S-0154018.
PK 16-SEP-1999: 990S-0154039.
PK 20-SEP-1999: 990S-0154779.
PK 22-SEP-1999: 990S-0155139.
PK 23-SEP-1999: 990S-0155486.
PK 24-SEP-1999: 990S-0155669.
PK 28-SEP-1999: 990S-0156458.
PK 29-SEP-1999: 990S-0156596.
PK 04-OCT-1999: 990S-0157117.
PK 05-OCT-1999: 990S-0157753.
PK 06-OCT-1999: 990S-0157865.
PK 07-OCT-1999: 990S-0158329.
PK 08-OCT-1999: 990S-0158332.
PK 12-OCT-1999: 990S-0158369.
PK 13-OCT-1999: 990S-0158929.
PK 13-OCT-1999: 990S-0159294.
PK 13-OCT-1999: 990S-0159295.
PK 14-OCT-1999: 990S-0159329.
PK 14-OCT-1999: 990S-0159330.
PK 14-OCT-1999: 990S-0159331.
PK 14-OCT-1999: 990S-0159637.
PK 14-OCT-1999: 990S-0159638.
PK 18-OCT-1999: 990S-0159584.
PK 21-OCT-1999: 990S-0160741.
PK 21-OCT-1999: 990S-0160767.
PK 21-OCT-1999: 990S-0160768.
PK 21-OCT-1999: 990S-0160770.
PK 21-OCT-1999: 990S-0160814.

```





```

UY 1228 gtgcgaatgcatcgcacaaataattgagaggttcaatatataagacatttcctc 1287
DB 602 www. .... 661
UY 1288 tgggtggagagaggtttcttgatttaattctgtgatacccaattcgttatttgg 1347
DB 662 www. .... 721
UY 1348 aagaaatgagaagttgacatgaattttaaaaaaataaaaaaataaaaaa 1407
DB 722 www. .... 781
UY 1408 atttcctgggcgcgc 1421
DB 782 www. .... 795

RESULT 8
AAF58257
ID AAF58257 standard: DNA: 936 BP.
XX
AC AAF58257:
XX
DI 24-APR-2001 (first entry)
XX
DB Oligonucleotide D1954.
XX
KW Electron-transfer group: ETM; mismatch; genotyping.
XX
KW gene expression; ss.
XX
OS Synthetic.
XX
PN WO200107665-A2.
XX
PD 01-FEB-2001
XX
PF 26-JUL-2000; 2000WO-US20476.
XX
PR 26-JUL-1999; 99US-0146795.
XX
PR 17-MAR-2000; 2000US-0190259.
XX
PA (CLIN-) CLINICAL MICRO SENSORS INC.
XX
PI time RM:
XX
DR WPI: 2001-159728/16.
XX
PT Nucleic acids containing electron-transfer group, useful as labels in
PT hybridization assays, e.g. for genotyping, allowing repeat analyses on
PT a single surface
XX
PS Example 6; Page 127; 159pp; English.
XX
CC The present invention relates to a composition comprising two nucleic
CC acids each containing an electron-transfer group (ETM) having
CC different redox potentials. The invention is used for electronic
CC detection of nucleic acids, especially of substitutions (mismatches)
CC and single-nucleotide polymorphisms, e.g. for genotyping.
CC monitoring gene expression.
XX
SQ Sequence 936 BP; 5 A; 142 C; 7 G; 6 T; 776 other;

Query Match 9.18; Score 130; DB 22; Length 936;
Best Local Similarity 1.38; Pval. No. 2.5e-20;
Matches 10; Conservative 492; Mismatches 292; Indels 0; Gaps 0;

UY 628 ttgaaagaagcctactatcattcattcaataagatttcaaatgtccta 687
DB 2 www. .... 61
UY 688 gatccagaagtgatcctaataatgtatgtatgtatgaactcgtgtgaagcaatgtt 747

```

```

DB 62 www. .... 121
UY 748 ctgctcctcagacagcagagctgttacttggaactatagctatgtccattgttg 807
DB 122 www. .... 181
UY 908 aattggttcacaggtgattgatgatgaataaatatgaacacctcaataacacagct 867
DB 182 www. .... 241
UY 868 atttgatacacttgaggaagtgaaagatactagtgaagggagcagcaltcaaatgt 927
DB 242 www. .... 301
UY 928 atcataagagggtttgacttgatagcgtagaatgcagagagaataataatggta 987
DB 302 www. .... 361
UY 988 gattggaagattaccaaagttgtcagggtttcagaggtataattccaacacagt 1047
DB 362 www. .... 421
UY 1048 ttgacctgaatcatgacaaatataatgacaaattcattacatgtgtattcagatt 1107
DB 422 www. .... 481
UY 1108 tgaagcaagctacgaagacacacagatcatcctagtcttccaagattgagat 1167
DB 482 www. .... 541
UY 1168 agtllttagtltgtgaadaactgtgtccctatcacatattgcacatagaggtt 1227
DB 542 www. .... 601
UY 1228 gtgcgaatgcatcgcacaaataattgagaggttcaatatataagacatttcctc 1287
DB 602 www. .... 661
UY 1288 tgggtggagagaggtttcttgatttaattctgtgatacccaattcgttatttgg 1347
DB 662 www. .... 721
UY 1348 aagaaatgagaagttgacatgaattttaaaaaaataaaaaaataaaaaa 1407
DB 722 www. .... 781
DB 782 www. .... 795

RESULT 9
AAF58259
ID AAF58259 standard: DNA: 936 BP.
XX
AC AAF58259:
XX
DI 24-APR-2001 (first entry)
XX
DB Oligonucleotide D2004.
XX
KW Electron-transfer group: ETM; mismatch; genotyping.
XX
KW gene expression; ss.
XX
OS Synthetic.
XX
PN WO200107665-A2.
XX
PI 01-FEB-2001.
XX
PR 26-JUL-2000; 2000WO-US20476.
XX

```





```

Db 122 www..... 181
QY 808 aattggttcaaggagatgagatgaatgaacattcaatataccagct 867
Db 182 www..... 241
QY 868 atttgaatgattgagagatgaagatagtgagagagagatgattgattg 927
Db 242 www..... 301
QY 928 atatatagagaggtttgtttgtatgagtgagtgagtgagtgagtgagtg 987
Db 302 www..... 361
QY 988 gaaggagaaattacaaaggttgcagggccttcacagactataatcaacaggt 1047
Db 362 www..... 421
QY 1048 ttgactgaataacaggaacataatagacaattcactcaattgtgattcagatt 1107
Db 422 www..... 481
QY 1108 tggagagaaagctacaggaagatgagatgagtgagtgagtgagtgagtg 1167
Db 482 www..... 541
QY 1168 agtttttttttttttttttttttttttttttttttttttttttttttttt 1227
Db 542 www..... 601
QY 1228 gtgcaatgttttgcaggaagatttgaggggtcgaatcttgagagatttgcct 1287
Db 602 www..... 661
QY 1288 tctgtgagagagagatgttttctgttttaaatctgtgttaacaaatcgtaatgttgg 1347
Db 662 www..... 721
QY 1348 aagaaatgagagattggaacttgaaatttttaaaaaaaagagagagagagag 1407
Db 722 www..... 781
QY 1408 attcctgagcgcg 1421
Db 782 wwwgagccccc 795

```

```

P1 Unk RM:
XX
XX DR WPI: 2001-159728/16.
XX
PT Nucleic acids containing electron-transfer group, useful as labels in
PT hybridization assays, e.g. for genotyping, allowing repeat analyses on
PT a single surface
XX
XX PS Example 6; Page 127; 159pp; English.
XX
CC The present invention relates to a composition comprising two nucleic
CC acids each containing an electron-transfer group (ETM) having
CC different redox potentials. The invention is used for electro-
CC detection of nucleic acids, especially of substitutions (mismatches)
CC and single-nucleotide polymorphisms, e.g. for genotyping,
CC monitoring gene expression.
XX
XX SO Sequence 938 BP; 4 A; 14 C; 9 G; 5 T; 776 other:

Query Match 9.1%; Score 100; DP 22; Length 938;
Best Local Similarity 1.3%; Prod. No. 2.5e-20;
Matches 149; Conserved 116; 40%; Mismatches 29%; Indels 0; Gaps 0;

QY 628 ttgaaagagagcctactatctccaatctcagaagattccatgttccatgcta 687
Db 2 www..... 61
QY 688 gatccaaagatgtgttccaaatgtgtatgtgttgaactcgtgtgaagcaatgt 747
Db 62 www..... 121
QY 748 ctgactcttgaatgagagagtggttttatttgggaattatgagatgacattgctg 807
Db 122 www..... 181
QY 808 aattggttcaaggagatgagatgaatgaacattcaatataccagct 867
Db 182 www..... 241
QY 868 atttgaatgattgagagatgaagatagtgaggggtcgaatcttgagagatttgcct 927
Db 242 www..... 301
QY 928 atatatagagaggtttgtttgtatgagtgagtgagtgagtgagtgagtgagtg 987
Db 302 www..... 361
QY 988 gaaggagaaattacaaaggttgcagggccttcacagactataatcaacaggt 1047
Db 362 www..... 421
QY 1048 ttgactgaataacaggaacataatagacaattcactcaattgtgattcagatt 1107
Db 422 www..... 481
QY 1108 tggagagaaagctacaggaagatgagatgagtgagtgagtgagtgagtgagtg 1167
Db 482 www..... 541
QY 1168 agtttttttttttttttttttttttttttttttttttttttttttttttt 1227
Db 542 www..... 601
QY 1228 gtgcaatgttttgcaggaagatttgaggggtcgaatcttgagagatttgcct 1287
Db 602 www..... 661
QY 1288 tctgtgagagagagatgttttctgttttaaatctgtgttaacaaatcgtaatgttgg 1347
Db 662 www..... 721
QY 1348 aagaaatgagagattggaacttgaaatttttaaaaaaaagagagagagagag 1407

```

```

RESULT 11
AAF58255
ID AAF58255 standard: DNA: 938 BP
XX
AC AAF58255:
XX
DT 24 APR 2001 (first entry)
XX
DE Oligonucleotide D1876.
XX
KM Election-transfer group, ETM, mismatch, genotyping;
KM gene expression; SS.
XX
OS Synthetic.
XX
PN W0200107565-A2.
XX
FE 01 FEB 2001
XX
PE 26-JUL-2000; 2000MO-JUL200476.
XX
PR 26-JUL-1999; 99MS-0145665
PR 17-MAR-2000; 2000MS-0190259
XX
XX (CLIN-) CLINICAL MICRO SENSORS INC.
XX

```

[illegible]

## RESULTS

ID AAF58252 standard; DNA; 936 bp.

AAH58252

24-APR-2001 (first entry)

DE oligonucleotide D1835.

KW Electron-transfer group; EFM mismatch; genotyping; homo dimerization; m

xxxxxx  
xxxxxxxx

XX  
XX  
WV0010766XX  
(1) - FEB-2001XX  
XX  
26-III-2001

XX 26-JUN-1999. 0911S-0145695  
 XX  
 PH

17-MAR-2000; 2000US-0190259;

PA (GAIN-) CLINICAL, MICRO SENS

Omega RM,  
XX  
XX

XX  
DN  
XX

01/07/2001, 17:00

17 Nucleic acids containing electron-transfer group, useful as labels in  
 18 hybridization assays, e.g. for genotyping, allowing repeat analyses on  
 19 a single surface -  
 20  
 21  
 22  
 23  
 24  
 25  
 26  
 27  
 28  
 29  
 30  
 31  
 32  
 33  
 34  
 35  
 36  
 37  
 38  
 39  
 40  
 41  
 42  
 43  
 44  
 45  
 46  
 47  
 48  
 49  
 50  
 51  
 52  
 53  
 54  
 55  
 56  
 57  
 58  
 59  
 60  
 61  
 62  
 63  
 64  
 65  
 66  
 67  
 68  
 69  
 70  
 71  
 72  
 73  
 74  
 75  
 76  
 77  
 78  
 79  
 80  
 81  
 82  
 83  
 84  
 85  
 86  
 87  
 88  
 89  
 90  
 91  
 92  
 93  
 94  
 95  
 96  
 97  
 98  
 99  
 100  
 101  
 102  
 103  
 104  
 105  
 106  
 107  
 108  
 109  
 110  
 111  
 112  
 113  
 114  
 115  
 116  
 117  
 118  
 119  
 120  
 121  
 122  
 123  
 124  
 125  
 126  
 127  
 128  
 129  
 130  
 131  
 132  
 133  
 134  
 135  
 136  
 137  
 138  
 139  
 140  
 141  
 142  
 143  
 144  
 145  
 146  
 147  
 148  
 149  
 150  
 151  
 152  
 153  
 154  
 155  
 156  
 157  
 158  
 159  
 160  
 161  
 162  
 163  
 164  
 165  
 166  
 167  
 168  
 169  
 170  
 171  
 172  
 173  
 174  
 175  
 176  
 177  
 178  
 179  
 180  
 181  
 182  
 183  
 184  
 185  
 186  
 187  
 188  
 189  
 190  
 191  
 192  
 193  
 194  
 195  
 196  
 197  
 198  
 199  
 200  
 201  
 202  
 203  
 204  
 205  
 206  
 207  
 208  
 209  
 210  
 211  
 212  
 213  
 214  
 215  
 216  
 217  
 218  
 219  
 220  
 221  
 222  
 223  
 224  
 225  
 226  
 227  
 228  
 229  
 230  
 231  
 232  
 233  
 234  
 235  
 236  
 237  
 238  
 239  
 240  
 241  
 242  
 243  
 244  
 245  
 246  
 247  
 248  
 249  
 250  
 251  
 252  
 253  
 254  
 255  
 256  
 257  
 258  
 259  
 260  
 261  
 262  
 263  
 264  
 265  
 266  
 267  
 268  
 269  
 270  
 271  
 272  
 273  
 274  
 275  
 276  
 277  
 278  
 279  
 280  
 281  
 282  
 283  
 284  
 285  
 286  
 287  
 288  
 289  
 290  
 291  
 292  
 293  
 294  
 295  
 296  
 297  
 298  
 299  
 300  
 301  
 302  
 303  
 304  
 305  
 306  
 307  
 308  
 309  
 310  
 311  
 312  
 313  
 314  
 315  
 316  
 317  
 318  
 319  
 320  
 321  
 322  
 323  
 324  
 325  
 326  
 327  
 328  
 329  
 330  
 331  
 332  
 333  
 334  
 335  
 336  
 337  
 338  
 339  
 340  
 341  
 342  
 343  
 344  
 345  
 346  
 347  
 348  
 349  
 350  
 351  
 352  
 353  
 354  
 355  
 356  
 357  
 358  
 359  
 360  
 361  
 362  
 363  
 364  
 365  
 366  
 367  
 368  
 369  
 370  
 371  
 372  
 373  
 374  
 375  
 376  
 377  
 378  
 379  
 380  
 381  
 382  
 383  
 384  
 385  
 386  
 387  
 388  
 389  
 390  
 391  
 392  
 393  
 394  
 395  
 396  
 397  
 398  
 399  
 400  
 401  
 402  
 403  
 404  
 405  
 406  
 407  
 408  
 409  
 410  
 411  
 412  
 413  
 414  
 415  
 416  
 417  
 418  
 419  
 420  
 421  
 422  
 423  
 424  
 425  
 426  
 427  
 428  
 429  
 430  
 431  
 432  
 433  
 434  
 435  
 436  
 437  
 438  
 439  
 440  
 441  
 442  
 443  
 444  
 445  
 446  
 447  
 448  
 449  
 450  
 451  
 452  
 453  
 454  
 455  
 456  
 457  
 458  
 459  
 460  
 461  
 462  
 463  
 464  
 465  
 466  
 467  
 468  
 469  
 470  
 471  
 472  
 473  
 474  
 475  
 476  
 477  
 478  
 479  
 480  
 481  
 482  
 483  
 484  
 485  
 486  
 487  
 488  
 489  
 490  
 491  
 492  
 493  
 494  
 495  
 496  
 497  
 498  
 499  
 500  
 501  
 502  
 503  
 504  
 505  
 506  
 507  
 508  
 509  
 510  
 511  
 512  
 513  
 514  
 515  
 516  
 517  
 518  
 519  
 520  
 521  
 522  
 523  
 524  
 525  
 526  
 527  
 528  
 529  
 530  
 531  
 532  
 533  
 534

The present invention relates to a composition comprising two nucleic acids each containing an electron-transfer group (ETG) having different redox potentials. The invention is used for electronic detection of nucleic acids, especially of substitutions (mismatches) and single-nucleotide polymorphisms, e.g. for genotyping, monitoring gene expression.

50 Sequence 936 BP; 4 A; 139 C; 10 G; 7 T; 776 other;

Query Match	9.08;	Score 129;	DB 22;	Length 936;
Best local Similarity	0.88;	pred. NO. 4.2e-20;		
Matches	6;	Conservative 484;	Mismatches 279;	Indels 0;
			Gaps	0

```

yy      642  tactttcttatttccgaagaattccaaatggtttccatgctagtccccaagatq 701
yb      779  wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 720
yy      702  gtccaaatggttgaatgatatactcgtgtagaacaattctcatccttcgaac 761
yb      719  wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 660
yy      762  atcgaacgttacttacttgggaatattgcatatgacattgctatgcttccag 821
yb      659  wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 600
yy      822  gatttgatgatgaatgaattagaaccttaataatccagatalltgctcctc 881
yb      599  wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 540

```

882 gaggaatgaaagatataatggaatgagaaatatttcaattatcatatataaagq 941  
 Db 539 ww 480  
 Oy 942 ttgatcttgaagcagaatgcagaatgataatgatttgaaggaagaattt 1001  
 Db 479 ww 420  
 Oy 1002 aaagatttgcagggcttcacagagatcatatttaaaacagtttgcattaa 1061  
 Db 419 ww 360  
 Oy 1062 atgacaaactatagcaaatctacatcacattgatttgcatttgaagaaagc 1121  
 Db 359 ww 300  
 Oy 1122 cagagagacacagatataatcattgatttgcagattgattgagatttga 1181  
 Db 299 ww 240  
 Oy 1182 gtgaataaactgttgcctatcacatatactacatgaagatttgcacattat 1241  
 Db 239 ww 180  
 Oy 1242 caaaagagattgcagagaggttcacatataagaaagatttgcatttgcgtg 1301  
 Db 179 ww 120  
 Oy 1302 atgtttcttcaattaaatcgttgaacccaattgtaaatttgaagaaatgaa 1361  
 Db 119 ww 60  
 Oy 1362 tgaacatgaaattttaaaaaaataaaaaaataaaaaaataatt 1410  
 Db 59 ww 11

## RESULT 13

11) AAF58254 standard; DNA; 936 bp.

AC: AAF58254

24-APR-2001 (first entry)

Oligonucleotide D1875

KW Electron-transfer group; EIM; mismatch; genotyping.

XX  
XX

XX  
DN 130300107665-23XX  
PD 01-FEB-2001

XX 36-III-3000-100000-11870476  
 PE

XX	92015.0145665
IX	92015.0145665
XX	92015.0145665

19. 1/-MAR-2000; 2000US-0190259-XX

PA (CLIN-) CLINICAL, MICRO SENSORS INC.  
XX XX

P1 Umex KM;  
XX

XX Example 6; Page 127; 159pp; English.





GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 31, 2001, 13:25:47 ; Search time 1922.14 Seconds  
(without alignments)  
11483.283 Million cell updates/sec

Title: US-09-577-657A-2

Sequence: 1427  
1 tggatcagctgtgggag . attttgggtttgggatttc 1427

Scoring table: IDENTITY\_NUC  
Gapop 10.0 ; Gapext 1.0

Searched: 1344157 seqs, 773874588 residues

Total number of hits satisfying chosen parameters: 2688314

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl: \*  
1: qb\_ba1: \*  
2: qb\_ba2: \*  
3: qb\_ba3: \*  
4: qb\_ba4: \*  
5: qb\_ba5: \*  
6: qb\_ba6: \*  
7: qb\_ba7: \*  
8: qb\_ba8: \*  
9: qb\_ba9: \*  
10: qb\_ba10: \*  
11: qb\_ba11: \*  
12: qb\_ba12: \*  
13: qb\_ba13: \*  
14: qb\_ba14: \*  
15: qb\_ba15: \*  
16: qb\_ba16: \*  
17: qb\_ba17: \*  
18: qb\_ba18: \*  
19: qb\_ba19: \*  
20: qb\_ba20: \*  
21: qb\_ba21: \*  
22: qb\_ba22: \*  
23: qb\_ba23: \*  
24: qb\_ba24: \*  
25: qb\_ba25: \*  
26: qb\_ba26: \*  
27: qb\_ba27: \*  
28: qb\_ba28: \*  
29: qb\_ba29: \*  
30: qb\_ba30: \*  
31: qb\_ba31: \*  
32: qb\_ba32: \*  
33: qb\_ba33: \*  
34: qb\_ba34: \*  
35: qb\_ba35: \*  
36: qb\_ba36: \*  
37: qb\_ba37: \*  
38: qb\_ba38: \*  
39: qb\_ba39: \*  
40: qb\_ba40: \*  
41: qb\_ba41: \*  
42: qb\_ba42: \*  
43: qb\_ba43: \*

44: em\_ov: \*  
45: em\_pat: \*  
46: em\_ph: \*  
47: em\_pl: \*  
48: em\_ro: \*  
49: em\_scs: \*  
50: em\_sy: \*  
51: em\_un: \*  
52: em\_v1: \*  
53: qb\_scs1: \*  
54: qb\_scs2: \*  
55: qb\_scs3: \*  
56: qb\_sy: \*  
57: qb\_un: \*  
58: qb\_v11: \*  
59: qb\_v12: \*  
60: qb\_hq1: \*  
61: qb\_hq2: \*  
62: qb\_hq3: \*  
63: qb\_hq4: \*  
64: qb\_hq5: \*  
65: qb\_hq6: \*  
66: qb\_hq7: \*  
67: qb\_hq8: \*  
68: qb\_hq9: \*  
69: qb\_hq10: \*  
70: qb\_hq11: \*  
71: qb\_hq12: \*  
72: qb\_hq13: \*  
73: qb\_hq14: \*  
74: qb\_hq15: \*  
75: qb\_hq16: \*  
76: qb\_hq17: \*  
77: qb\_hq18: \*  
78: qb\_hq19: \*  
79: qb\_hq20: \*  
80: qb\_hq21: \*  
81: qb\_hq22: \*  
82: qb\_hq23: \*  
83: qb\_hq24: \*  
84: qb\_hq25: \*  
85: qb\_pr1: \*  
86: qb\_pr2: \*  
87: qb\_pr3: \*  
88: qb\_pr4: \*  
89: qb\_pr5: \*  
90: qb\_pr6: \*  
91: qb\_pr7: \*  
92: qb\_pr8: \*  
93: qb\_pr9: \*  
94: qb\_pr10: \*  
95: qb\_pr11: \*  
96: qb\_pr12: \*  
97: qb\_pr13: \*  
98: qb\_pr14: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1403.2	98.4	1408	12	AB041280
2	239.2	16.7	1447	15	AB041280
3	215.4	15.1	1152	15	AF0308570
4	213.8	15.0	1321	12	AF133053
5	155.6	10.9	1306	13	AF179232
6	151.6	10.6	1404	12	AB044792
7	151.2	10.6	1316	12	AB048793
8	149.2	10.5	1359	12	AB039725

9	138.8	9.7	139.8	12	AB048794	Coffea arabica
10	134.4	9.4	136.3	13	AF198492	AFL198492 Arabidopsis
11	103.8	7.3	93890	12	AC006528	AB006528 Arabidopsis
12	101.4	7.1	82348	12	AB028622	AB028622 Arabidopsis
13	100.6	7.0	61510	12	AB028606	AB028606 Arabidopsis
14	93.2	6.5	198354	13	ATAP22	779708 Arabidopsis
15	93.2	6.5	198750	13	ATACHRV85	AB116589 Arabidopsis
16	92.8	6.5	83511	12	AB013389	AB013389 Arabidopsis
17	81	5.9	117206	12	AC008915	AC008915 Arabidopsis
18	83	5.8	90341	12	AC009460	AC009460 Arabidopsis
19	79.6	5.6	454	14	HVC234779	AJ231779 Arabidopsis
20	71	5.0	161565	12	AC012563	AC012563 Arabidopsis
21	69.6	4.9	149477	83	AP003378	AP003378 Arabidopsis
22	69.6	4.9	160468	83	AP003245	AP003245 Arabidopsis
23	69.2	4.8	154679	83	AP003274	AP003274 Arabidopsis
24	69	4.8	49890	12	AF028010	AF028010 Arabidopsis
25	55.8	3.9	78239	13	AB011241	AB011241 Arabidopsis
26	54.2	3.8	1084	53	CNS06PBLB	AL005097 end of
27	52.6	3.7	2209	53	AF180728	AL180728 Arabidopsis
28	52.6	3.7	2550	93	HSM801250	AF127045 Homo sapi
29	52.6	3.7	124006	87	AF068882	AC008882 Homo sapi
30	52.6	3.7	145264	87	AC016601	AC016601 Homo sapi
31	52.6	3.7	160584	69	AP026006	AC026006 Homo sapi
32	52.6	3.7	164399	96	PFMAL336	298551 Plasmodium
33	52.4	3.7	2475	95	HVC65274	066274 Rattus norv
34	52.4	3.7	2481	9	AF079011	AB079011 Sequence
35	52.4	3.7	2604	97	AP079012	AP079012 Scyphozoa
36	52	3.6	141984	14	FNLI	AC005551 Arabidops
37	51.4	3.6	1180	91	BC005379	BC005379 Homo sapi
38	51.4	3.6	49850	13	ATF28D10	AL131254 Arabidops
39	51	3.6	759	53	CNS06OXV	AL111257 end of
40	51	3.6	47460	12	AB011478	AB011478 Arabidops
41	51	3.6	164445	60	AC008742	AC008742 Homo sapi
42	51	3.6	214094	77	AC0087188	AC0087188 Homo sapi
43	50.4	3.5	1868	89	AF216052	AF216052 Homo sapi
44	50.2	3.5	2192	55	AY008296	AC008296 Tribolium
45	50.2	3.5	143172	70	AP026462	AC026462 Homo sapi

## ALIGNMENTS

90. 1199  
/gene="TCS1"  
/function="N-methyltransferase"  
/codon\_start=1  
/product="cat;cathe synthase"  
/protein\_id="BAB12278.1"  
/db\_xref="GI:9667143"  
/translation="MELIAFARVAVLEPRNKEHSSVYANSSFTQGVASMAQPALE  
AVETLSKDFHGLANAAHLEAAAPNIPAVISTIKMKMKRKRLENTVLTLOVLYL  
L PGNPNPTLFFPSSSEVIGNKTETVEPVYVYVSEHGLIPRNSLIVHSSVYVH  
LIAPGKGLSKRGLALNNGKIVTSKISHPVAVKAVLSQPHEDTTLNRSQVAPN  
CVLLIKGKGNSTPSIDMOSCTLWELIMAAVLAVSQGLDHRKIDFENIPSVASTLE  
VKRIYERDGSFTLIDIEGPDLSVEMQENKRWKGEFTKVVAATFPIISNORPE  
MKRLTRFHIVSELEAKLEPFTISLTSKIDG"

blast -c:cdnt 453 -s 248 -q 426 -g 411 -t  
ORIGIN

Query Match: 92.36, Score 145.27, E=1.2, Length 1438;  
Posit Local Similarity: 99.84, Ident No. 9,556/241,  
Matches 1405; Conservative 0; Mismatches 3; Indels 0; Gaps 0.

QY 1 Tgatactacgcgcgtgcgcgcgcgcgcctcttctctatataaaatctcttccgcgaagca 60  
RQ 31 TTAAT 90  
QY 61 TGGAGCTAGCTACG 120  
DB 91 GCGAGCTAGCTAT 150  
QY 121 AAGAGTAT 180  
DB 151 AAGAGTAT 210  
QY 181 AAGAGTAT 240  
DB 211 AAGAGTAT 270

TACUS AB031280 1438 bp mRNA PIN 31-AUG-2000  
 DEFINITION *Camellia sinensis* TCST1 mRNA for caffeine synthase, complete cds.  
 ACCESSION AB031280  
 VERSION AB031280.1 GI:99677142  
 KEYWORDS caffeine synthase.  
 SOURCE *Camellia sinensis* cDNA to mRNA.  
 ORGANISM *Camellia sinensis*  
 Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta  
 Magnoliophyta; Eudicotyledons; Core eudicot; Ascertidae; Fricales;  
 Theaceae; *Camellia*.  
 1 (sites)  
 Kato, M., Mizuno, K., Crozier, A., Fujimura, T. and Ashihara, H.  
 Caffeine synthase gene from tea leaves  
 Nature 406 (6799), 956-957 (2000)  
 20437335  
 2 (bases 1 to 1438)  
 Kato, M., Mizuno, K., Ashihara, H. and Fujimura, T.  
 Direct Submition  
 Submitted (16-AUG-1999) to the EMBL/GenBank and DDBJ  
 Mizuno, University of Tsukuba, Institute of Agricultural and Forest  
 Engineering, 1-1-1 Tennoudai, Tsukuba, Ibaraki 305-8572, Japan  
 (E-mail: komomo@sakura.cc.tsukuba.ac.jp, Tel: 81-298-53-4656,  
 Fax: 81-298-55-2203)  
 FEATURES  
 SOURCE  
 Location/Qualifiers  
 1. 1438  
 /organism="Camellia sinensis"  
 /db\_xref="taxon:4442"  
 90..1399  
 /gene="TCST1"

Query Matched: 92.36, Score: 106.97, Db: 12, Length: 1438;  
Post Local Similarity: 99.88, E-Val: No, 5.5e-241  
Matches: 1405, Conservative: 0, Mismatches: 3, Indels: 0, Gaps: 0

```
QY      1  lqataactgctgtggaagctgcctcttgcgtataaaatctcttctgcagaaga 60
RB      31  TATAATCAATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 90
QY      61  tggatctgagctactggaggggaaggggaaggtttgctctctctctctgggggga 120
DB      91  TGGAGAGTACGCTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 150
QY      121  aaaggttatatgacaaactctctctctctctctctctctctctctctctctct 180
RB      151  AAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 210
QY      181  agctaaagatgaggttgaagctctctctctctctctctctctctctctctctct 240
RL      211  GATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 270
QY      241  cagaggaactggaggttgcgcagagagctcgaagaacttcgaatattctacatt 300
DB      271  CAGGCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 330
QY      301  gaatgatggaagaagaatgcaggaatctgaattgcgaagaactggaacttcgaat 360
DB      331  GATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 390
QY      361  tgaatgctctcttcttggaaatgatactcaataacctctcgaagccctcgtctaa 420
DB      391  TGAATGATGCTCTTTTGGAGAAAGATTTCATTAATACCTCTTTTAAAGAGGCT 450
QY      421  ttgtaacaatatgtaaggaagttcctgtctgtatgtaggaatcgcaggtcttct 480
DB      451  TTGCTTAATTAATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 510
QY      481  ggcgcctcttctctcgtaaagcttaacatttaatttattctctctaaagttcatt 540
DB      511  GCGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 570
QY      541  ttaactggagcaacaaagagactcagaatgaggaagctcttgctctaaagaaga 600
DB      571  TTACTCGAGGACCAAAAGACACTCAAAAGACGAGAGGCTTGCTATTAAACAAAG 630
QY      601  cttaactgcaacaaagagagagctctgtctatgaggaagctctctctctctct 660
DB      631  CTAACTGCAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAA 690
QY      661  aagatctcaaatgctctctgaatctatgatacgaagagaggtctctgaatgata 720
DB      691  AAGATTCACAAATGTTCTCTCAATCTCAATCAATCAATCAATCAATCAATCAAT 750
QY      721  tctgtaactctctgttagcaaatctctaaactcttgagatcgcaagatgcttact 780
DB      751  TCTGTGAATCTTCTGTGTGAATCTCTGTGTGAATCTTGAATCTTGAATCTTGA 810
```

QY 781 ggaactatctagctatgagccatctgcaacttgcttcaaggatctgagtgaagata 840  
 |||||  
 Db 811 GGGACACATATAGCTATGGGCTGCTGCAATGCTGTTTACAGGCGATGATAGAGATA 870  
 |||||  
 QY 841 aattagacccttcaataaccagcattctgacacttgaggaatgaagaatata 900  
 |||||  
 Db 871 AATTAGACATCTTAAATATATCCAGTATTATGATATACCTTGAGGAAATGATATAG 930  
 |||||  
 QY 901 tggagggagcagatctctcaaatctgacatctgaggggttctgactctgatacga 960  
 |||||  
 Db 931 TGAGAGAGAGAGATATATGATATGATATGATATGATATGATATGATATGATATG 990  
 |||||  
 QY 961 gaatgagcagatctctcaaatctgacatctgaggggttctgactctgatacga 1020  
 |||||  
 Db 991 AATATGATATGATATGATATGATATGATATGATATGATATGATATGATATGATATG 1050  
 |||||  
 QY 1021 tcaagagcagatctctcaaatctgacatctgaggggttctgactctgatacga 1080  
 |||||  
 Db 1051 TCAGAGAGAGAGATATGATATGATATGATATGATATGATATGATATGATATGATATG 1110  
 |||||  
 QY 1081 aattgagcagatctctcaaatctgacatctgaggggttctgactctgatacga 1140  
 |||||  
 Db 1111 AATATGATATGATATGATATGATATGATATGATATGATATGATATGATATGATATG 1170  
 |||||  
 QY 1141 tcaagagcagatctctcaaatctgacatctgaggggttctgactctgatacga 1200  
 |||||  
 Db 1171 TCAAGAGAGAGATATGATATGATATGATATGATATGATATGATATGATATGATATG 1230  
 |||||  
 QY 1201 tcaagagcagatctctcaaatctgacatctgaggggttctgactctgatacga 1260  
 |||||  
 Db 1231 TCAGAGAGAGAGATATGATATGATATGATATGATATGATATGATATGATATGATATG 1290  
 |||||  
 QY 1261 ggaactatctagctatgagccatctgcaacttgcttcaaggatctgagtgaagata 1320  
 |||||  
 Db 1291 GGGACACATATAGCTATGGGCTGCTGCAATGCTGTTTACAGGCGATGATAGAGATA 1350  
 |||||  
 QY 1321 ctgagatcccaaatctgacatctgaggggttctgactctgatacga 1380  
 |||||  
 Db 1351 CTGATGATATGATATGATATGATATGATATGATATGATATGATATGATATGATATG 1410  
 |||||  
 QY 1381 aattgagcagatctctcaaatctgacatctgaggggttctgactctgatacga 1408  
 |||||  
 Db 1411 AATATGATATGATATGATATGATATGATATGATATGATATGATATGATATGATATG 1438  
 |||||

RESULT 2  
 AB049752 1437 bp mRNA PLN 16-MAR-2001  
 LOCUS  
 DEFINITION Atropa belladonna AbsAMT1 mRNA for S-adenosyl-L-methionine:salicylic acid carboxyl methyltransferase, complete cds  
 ACCESSION AB049752  
 VERSION AB049752.1 GI:13366160  
 KEYWORDS  
 SOURCE Atropa belladonna (strain M8) root cDNA for mRNA, clone lambda ABSAMT1.  
 ORGANISM Atropa belladonna  
 Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; Asterids; Ericales; Solanales; Solanaceae; Atropa.  
 REFERENCE 1 (sites)  
 AUTHORS Fukami, H., Asakura, T., Hirano, H., Abe, K., Shimomura, K. and Yamakawa, T.  
 TITLE Cloning and expression of salicylic acid inducible and active S-adenosyl-L-methionine:salicylic acid carboxyl methyltransferase in transformed root culture of Atropa belladonna  
 JOURNAL unpublished  
 REFERENCE 2 (bases 1 to 1437)  
 AUTHORS Fukami, H., Asakura, T., Hirano, H., Abe, K., Shimomura, K. and Yamakawa, T.  
 TITLE Direct Submission  
 JOURNAL Submitted (10-OCT-2000) Takashi Yamakawa, the University of Tokyo,

Department of Global Agricultural Sciences: 1-1, Yayoi 1-chome, Bunkyo-ku, Tokyo 113-8657, Japan  
 E-mail: ayama@mail.ecc.u-tokyo.ac.jp, Tel: 81-3-5841-7515, Fax: 81-3-5841-5304

FEATURES  
 source  
 location/Qualifiers  
 1. 1437  
 /organism="Atropa belladonna"  
 /strain="M8"  
 /db\_xref="taxon:33113"  
 /clone="lambda AbsAMT1"  
 /tissue\_type="root"  
 62. 1135  
 /gene="ABSAMT1"  
 62. 1135  
 /gene="ABSAMT1"  
 /odon\_start=1  
 /product="S-adenosyl-L-methionine:salicylic acid carboxyl methyltransferase"  
 /protein\_id="BAB39396.1"  
 /db\_xref="GI:13366161"  
 /translation="MKVVEVLMHNGNDISANNISIVKRVILMTKPTEDQISDLY  
 CSFEPLTCLALDSSSANTLVELEVALEKRLKLNLSANLHFNLDLGNDF  
 NLFQSLKRFQDNIKRLQSEFGDFFSVSPSTPTPFPSSTHFPVSSYIMLSQ  
 VPLLRKKNENIYIASISPEVYIKVYKQVKKDSNLIKPSFELMGKKVLLFELR  
 ESDDPSKRCQYIMELSMALNELVLELIIEKKDPSNIPQVTPSPPEVYIVKES  
 SEIINLEIIVPHKVNINQINDVNVAKCHRAVAGFLVLSLTLQKLMNVFYCYEEL  
 ISDLSKEETFTNTVSLTKR"

FAST COUNT 488 a 225 c 290 g 434 t  
 ORIGIN  
 Query Match 14.7% Score 498.2, DR 123, Length 1437  
 Best Local Similarity 54.98% Pred. No. 3,8e-41  
 Matches 599, Conservative 0, Mismatches 453, Indels 39, Gaps 5,  
 QY 80 gaagtgagcagatctctcaaatctgacatctgaggggttctgactctgatacga 139  
 |||||  
 Db 64 GAAGTGATATGATATGATATGATATGATATGATATGATATGATATGATATGATATGATATG 123  
 |||||  
 QY 140 ctctctcttcaagcagaaagctgacatctgaggggttctgactctgatacga 199  
 |||||  
 Db 124 CTCTCTCTTCAAGCAGAAAGCTGACATCTGAGGAGAAAGCTGAGGAGGAGGAGGAGGAG 183  
 |||||  
 QY 200 ggaactatctagctatgagccatctgcaacttgcttcaaggatctgagtgaagata 259  
 |||||  
 Db 184 GGAACATATAGCTATGGGCTGCTGCAATGCTGTTTACAGGCGATGATAGAGATA 240  
 |||||  
 QY 260 aacgagtcac 319  
 |||||  
 Db 241 ATTCAGAGAGAGAGATATGATATGATATGATATGATATGATATGATATGATATGATATG 300  
 |||||  
 QY 320 ggaagatctgagatctgac 379  
 |||||  
 Db 301 AAAAAATATATATTTTAAAT 360  
 |||||  
 QY 380 tgaatgagcagatctctcaaatctgacatctgaggggttctgactctgatacga 432  
 |||||  
 Db 361 TGAATGATATGATATGATATGATATGATATGATATGATATGATATGATATGATATGATAT 420  
 |||||  
 QY 433 tgaatgagcagatctctcaaatctgacatctgaggggttctgactctgatacga 487  
 |||||  
 Db 421 TGAGAGAGAGAGATATGATATGATATGATATGATATGATATGATATGATATGATATGATAT 480  
 |||||  
 QY 488 ttctctcgaagcagatctgagatctctctctctctctctctctctctctctctctctct 547  
 |||||  
 Db 461 TTCTCTCGAAGCAGATCTGAGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 540  
 |||||  
 QY 548 ggaacacaaagac 607  
 |||||  
 Db 541 AGTTCGATTTAATTAATAAG-----AACAGAGAGAAATTTTAT 592  
 |||||  
 QY 608 atcaagacagccctctgctgtaagagagacacacacacacacacacacacacacacacacac 667  
 |||||





[illegible][illegible]





[illegible][illegible]











```

CD5 /note="contains similarity to receptor protein kinase
gene_id:MEN24.4"
/codon_start=1
/evidence=not_experimental
/protein_id="BAB01369.1"
/db_xref="GI:9279743"
/translation="MSSSSVSKRPVPIVIVVTTQLLIPVNSLNTNSYLHKKK
VWVGQYKPSKVEKSLDIQSNNKIKQSYSPPIQSKAKGKPKWVSITYGCE
DSGPGKQSVVATGATGELLKKCPYKAKLITWQGLVESSLDSGQINDNDNMCM
SANLIONSLSLEERHLNNITKATIKDKINBGIKKVYLAAGEKDLQSKLQK
VGSANLSVGRNFCMIYTVHPEDTWKQGVHVSFSPSNPEYELYPINRPPYTA
K"
join(14194..14550,14764..15243)
/note="contains similarity to receptor protein kinase
gene_id:MEN24.5"
/codon_start=1
/feature="not_expected"
/protein_id="BAB01370.1"
/db_xref="GI:9279744"
/translation="MSSSSVSKRPVPIVIVVTTQLLIPVNSLNTNSYLHKKK
VWVGQYKPSKVEKSLDIQSNNKIKQSYSPPIQSKAKGKPKWVSITYGCE
DSGPGKQSVVATGATGELLKKCPYKAKLITWQGLVESSLDSGQINDNDNMCM
SANLIONSLSLEERHLNNITKATIKDKINBGIKKVYLAAGEKDLQSKLQK
VGSANLSVGRNFCMIYTVHPEDTWKQGVHVSFSPSNPEYELYPINRPPYTA
K"
join(14194..14550,14764..15243)
CD6 /note="contains similarity to receptor protein kinase
gene_id:MEN24.6
pit1p110607
similar to unknown protein"
/codon_start=1
/evidence=not_experimental
/protein_id="BAB01371.1"
/db_xref="GI:9279745"
/translation="MPPPPSPKRIYVHISAVVAIKLLIKSSSLNNMNPVLDH
CIYSEKQYGVGEYENKINILIPASSAIDFESDGVVPSKSNENILMYQPEKSY
SPRSPYTAIAAAYVAMGFHVSIVKIKYAAVWV"IKHIMIKIYALFEMITKSP
CDKLGKRGVNSCRNRELYPIKPY"
join(17846..17907,18449..18770,18904..19402)
/note="contains similarity to receptor protein kinase
gene_id:MEN24.7"
/codon_start=1
/evidence=not_experimental
/protein_id="BAB01372.1"
/db_xref="GI:9279746"
/translation="MHANSGIDHCKITPLPYLGSGIALAIPNVLSLNGINAVY
HHQWQDQGYNSYTPYKRNINSILPTISINIPGVHSSNDANSVFKIGPRT
SYSKSCSLQATANSIEPRPKNSKEKLIWQDNCLELISIDFPKIDYQNNFMYN
KDSYDITFEFNEFTLHLKFNANLSKQIPYAAVAFERILSKIKYAAVW"IKHIE
PNNQNVLWIMAKLIPCKQKQGVSTSCNRELYPELMI"
join(20711..21066,21605..22076)
/note="contains similarity to receptor protein kinase
gene_id:MEN24.8"
/codon_start=1
/feature="not_expected"
/protein_id="BAB01373.1"
/db_xref="GI:9279747"
/translation="MSSSSASKLQSVIVFAMISVQVLFTHCWMSLNTNSNOTNV
LHHTINSQGYKPSQSPYENINRYATISITTSNLSGIDTEYFNYNKKDQ
TITQPSAASPLSPDPRCKNKGRIWQDNCLELISIDFPKIDYQNNFMYNKKDQ
NGQINFEINPALLYFLKRAKISENNNRQVPLISNIDAKMIYATKSTFIM
KLGVQWQYKIDSKITVCLNMLIILAKVITQSSKMSGVPYSTSCNRELYPELMI"
join(23168..23494,23739..24185)
/note="gblAAD24839.1
gene_id:MEN24.9
similar to unknown protein"
/codon_start=1
/evidence=not_experimental
/protein_id="BAB01374.1"
/db_xref="GI:9279748"
/translation="MSSSPYKSLCTVITAAVAMQIFQSVSLQCTREYIHTGCT
NRQGYSEKQVARDLNDLQMSISLIGDVYSGNSIYVLDQPDGSSITQSRSL
LTAFSQEMSGKAPNNVYQCLSHNPISTVWFCRLPQFTFNKQVSTAAFSQSL
KITRLIKLKEKASSKKNIPKKNITVATGESGAKKILAMVQCTQDLSKNQVCI

```

[illegible]





The annotation of this entry was produced with considerable contributions from Stephane Rombauts and Pierre Rouze, Department of Genetics, University of Ghent, Leuvenstraat 35, 9000 Ghent, BE. E-mail: stromgen@ugr.ac.be, p.rouze@ugr.ac.be. Information on performance of analysis and a more detailed annotation of this entry and other sequences of chromosomes 3, 4 and 5 can be viewed at: <http://www.mips.biochem.mpg.de/proj/thal/> this fragment has an overlap with ATAP21 at the 5' end.

## FEATURES

SOURCE  
1. J98354  
/organism="Arabidopsis thaliana"  
/variety="Columbia"  
/db\_xref="taxon:3702"  
/chromosome="4"  
1. 10046  
/organism="Arabidopsis thaliana"  
/db\_xref="taxon:3702"  
/clone="cosmid c1c7A10.4F12"  
8165..10851  
/gene="C7A10.500"  
8165..8344  
/gene="C7A10.500"  
/number=1  
join(8165..8344,8417..8522,8639..8983,9116..9209,9502..9631,9729..9932,10023..10889,11065..10284,10372..10430,10523..10851)  
/gene="C7A10.500"  
note="Similar to hypothetical protein 1517 from Arabidopsis thaliana chromosome II BAC T517 genomic sequence, PATX:02642165  
Contains Neutral zinc metalloproteinases, zinc binding region signature [1:AAHEMMHMM]  
contains EST gb:N65197, H76120, AA06496"  
/codon\_start=1  
/product="putative Protein"  
/protein\_id="PARI6816.1"  
/db\_xref="GI:4006886"  
/db\_xref="SPTREMBL:O23197"  
/translation="MPFLIPLCTITVPIYAGKRIILPWKTKTIKGSNHSFGDQNGHYPIPIVFGPPYSAEGSDKEIEFGAIALSSEGHVLPQDQKRIIEKSPTEEDDDDEDEEYKRAOLEAAEEERVAQAOIEEEKRAEAOLEFEKILAKARLEHEMRKSAOLFEDELAKALOESNVGSPRPYDGNILQYPLIIPSSHRCYCGCAEIGHGRFLSAWAFGILNVSAAHPIIPNPAQILEKAPITWQKTCPSHEDGIPRCSEKMEPKUTKTLILDDGRKICLRLTUSAIMDHPQQLYLEIKELVSGIAMVQSNVAFQGEIISKSPAMPEKRGHHLPEPTGCLSEGTIVTLRPPIGACGLKIDMTFPCRLIRQCFVATILLYGLPRLTGLSTLAHEMMHMMALRLNGYNLRPEVEGICQVLAHMELETFYAGSTLVDIASSSSAVVASSKSGPSPDEKKLGEFFKHQESDSSA YDNGFQKNGAVLKHGLRRLIDHILRLIGTFPKKI"  
8145..8416  
/gene="C7A10.500"  
/number=1  
8417..8522  
/gene="C7A10.500"  
/number=2  
8523..8637  
/gene="C7A10.500"  
/number=2  
8638..8983  
/gene="C7A10.500"  
/number=3  
8984..9115  
/gene="C7A10.500"  
/number=4  
9116..9209  
/gene="C7A10.500"  
/number=4  
9210..9501  
/gene="C7A10.500"  
/number=4  
9502..9631  
/gene="C7A10.500"  
/number=5  
9632..9719

/gene="C7A10.500"  
/number=5  
9720..9932  
/gene="C7A10.500"  
/number=6  
9933..10026  
/gene="C7A10.500"  
/number=6  
10027..10089  
/gene="C7A10.500"  
/number=7  
10090..10165  
/gene="C7A10.500"  
/number=7  
10166..10284  
/gene="C7A10.500"  
/number=8  
10285..10371  
/gene="C7A10.500"  
/number=8  
10372..10430  
/gene="C7A10.500"  
/number=9  
10431..10522  
/gene="C7A10.500"  
/number=9  
10523..10851  
/gene="C7A10.500"  
/number=10  
13598..13824  
/gene="C7A10.510"  
/number=1  
13598..16004  
/gene="C7A10.510"  
join(13598..13824,13914..13956,14045..14071,14245..14343,14416..14582,14855..14999,15193..15234,15252..15417,15513..15533,15643..15792,15894..16004)  
/gene="C7A10.510"  
note="Intron number 9 is a special 912 intron similarity to various predicted proteins of Arabidopsis and yeast  
Contains Somatostatin, prolactin and related hormones signatures [CVALDLEFLLQYIKKVC]  
contains EST gb:H77118, AA065507"  
/codon\_start=1  
/product="putative protein"  
/protein\_id="CAH16817.1"  
/db\_xref="GI:4006887"  
/db\_xref="SPTREMBL:O23198"  
/translation="WVSLGVCIKKKTCVWPVLYDNCIQLNINIVSFAIGASLLGMCVAEIPQVITNERTKSSNVSISLLAWAGDIFNLVGLLEPAILPQFIYALLYT VSTVLVLIQTIYDLYKLCFHFPTKIQKDEDEEEKPLKPKTKMSAISIPGSKYDSSPEFYTSAPSLAGTPPLPTSEVPKASGLPIWQSTNMVPHPTTPQKRGAGCTFLAASASIPQAKSLAEKAYASAKRLINERIVRSALQQLGWI MAIYMGRIIP QIWNIRKQSVPGIIPLEMTALVANNITVGSILVPTTDNIRKPNILWLLIATVGV LHLITLIQYIKKVCIKSLPSKELAVGVASKTVFS"  
13825..13913  
/gene="C7A10.510"  
/number=1  
13914..13956  
/gene="C7A10.510"  
/number=2  
13957..14044  
/gene="C7A10.510"  
/number=2  
14045..14071  
/gene="C7A10.510"  
/number=3  
14072..14244  
/gene="C7A10.510"  
/number=3  
14245..14343  
/gene="C7A10.510"



exon	64.47..6553.3 /gene="AT4436190" /number=5	64.47..6553.3 /gene="AT4436190" /number=5
intron	6554..6644 /gene="AT4436190" /number=5	6554..6644 /gene="AT4436190" /number=5
exon	6645..6698 /gene="AT4436190" /number=6	6645..6698 /gene="AT4436190" /number=6
intron	6699..6751 /gene="AT4436190" /number=6	6699..6751 /gene="AT4436190" /number=6
exon	6752..6848 /gene="AT4436190" /number=7	6752..6848 /gene="AT4436190" /number=7
intron	6849..6965 /gene="AT4436190" /number=8	6849..6965 /gene="AT4436190" /number=8
exon	6966..7171 /gene="AT4436190" /number=8	6966..7171 /gene="AT4436190" /number=8
intron	7172..7264 /gene="AT4436190" /number=8	7172..7264 /gene="AT4436190" /number=8
exon	7265..7359 /gene="AT4436190" /number=9	7265..7359 /gene="AT4436190" /number=9
intron	7360..7417 /gene="AT4436190" /number=9	7360..7417 /gene="AT4436190" /number=9
exon	7418..7525 /gene="AT4436190" /number=10	7418..7525 /gene="AT4436190" /number=10
intron	7526..7652 /gene="AT4436190" /number=10	7526..7652 /gene="AT4436190" /number=10
exon	7653..7746 /gene="AT4436190" /number=11	7653..7746 /gene="AT4436190" /number=11
intron	7747..7853 /gene="AT4436190" /number=11	7747..7853 /gene="AT4436190" /number=11
exon	7854..7894 /gene="AT4436190" /number=12	7854..7894 /gene="AT4436190" /number=12
intron	7895..8795 /gene="AT4436190" /number=12	7895..8795 /gene="AT4436190" /number=12
exon	8796..8921 /gene="AT4436190" /number=13	8796..8921 /gene="AT4436190" /number=13
intron	8922..9011 /gene="AT4436190" /number=13	8922..9011 /gene="AT4436190" /number=13
exon	9012..9185 /gene="AT4436190" /number=14	9012..9185 /gene="AT4436190" /number=14
intron	9186..9264 /gene="AT4436190" /number=14	9186..9264 /gene="AT4436190" /number=14
exon	9265..9450 /gene="AT4436190" /number=15	9265..9450 /gene="AT4436190" /number=15
intron	9451..9551 /gene="AT4436190" /number=15	9451..9551 /gene="AT4436190" /number=15
exon	9552..9668 /gene="AT4436190" /number=16	9552..9668 /gene="AT4436190" /number=16
intron	9669..9767 /gene="AT4436190" /number=16	9669..9767 /gene="AT4436190" /number=16
exon	9768..9821	9768..9821

GenCore version 4.5  
Copyright (c) 1993 - 2000 GenCore Inc

OM protein - protein search, using sw model

Run on: August 31, 2001, 15:00:32 : Search time 59.12 Seconds

(without alignments)

796,655 Million cell updates/sec

File: us-09-577-657a-1

Perfect score: 1847

Sequence: 1 PNNRPPSSSYAANSSTFQ2

EEAVLTITIVSTIR 366

Scoring table:

Rossm62  
Gapop 10.0 : Gapext 0.5

Searched: 425026 seqs, 132305027 residues

Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 6

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: SP.TREMBL\_16:\*  
2: sp.archaea:\*  
3: sp.bacteria:\*  
4: sp.fungi:\*  
5: sp.invertebrate:\*  
6: sp.mammal:\*  
7: sp.mhc:\*  
8: sp.oraneller:\*  
9: sp.phage:\*  
10: sp.plant:\*  
11: sp.podent:\*  
12: sp.pneustic:\*  
13: sp.vertebrate:\*  
14: sp.virus:\*

Fred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No	Score	Query Match	Length	Fr	IT	Accession
1	1847	100.0	369	10	09FZNB	Q9FZNB camellia si
2	724	39.2	332	10	Q9SBK6	Q9SBK6 brassica ra
3	713	38.6	359	10	Q9SPV4	Q9SPV4 arabidopsi
4	637.5	34.5	354	10	Q9F176	Q9F176 arabidopsi
5	630	34.1	371	10	Q9Z234	Q9Z234 arabidopsi
6	621.5	33.9	364	10	Q9FV79	Q9FV79 arabidopsi
7	585.5	31.7	363	10	Q9LS20	Q9LS20 arabidopsi
8	582.5	31.5	368	10	Q9LRT5	Q9LRT5 arabidopsi
9	570	30.4	359	10	Q9ZPT3	Q9ZPT3 arabidopsi
10	499.5	27.0	351	10	Q9K157	Q9K157 arabidopsi
11	496.5	26.9	348	10	Q9HXC4	Q9HXC4 arabidopsi
12	485.5	26.3	448	10	Q9FV76	Q9FV76 arabidopsi
13	482.5	26.1	359	10	Q9LS10	Q9LS10 arabidopsi
14	481	26.0	386	10	Q9FLN8	Q9FLN8 arabidopsi
15	480	26.0	379	10	Q9FY73	Q9FY73 arabidopsi
16	480	26.0	379	10	Q9FY73	Q9FY73 arabidopsi
17	480	26.0	379	10	Q9FY73	Q9FY73 arabidopsi
18	475.5	25.7	353	10	Q9FWP6	Q9FWP6 arabidopsi
19	475.5	25.7	361	10	Q9FKR0	Q9FKR0 arabidopsi

20	466.5	25.3	360	10	Q9FW11	Q9FW11 oryza sativ
21	454	24.6	362	10	Q9FKD0	Q9FKD0 arabidopsi
22	453.5	24.5	389	10	Q91MA2	Q91MA2 arabidopsi
23	424	23.0	352	10	Q9M4F2	Q9M4F2 brassica na
24	387	21.0	318	10	Q9F113	Q9F113 brassica na
25	105.5	5.7	100	10	Q9LS11	Q9LS11 arabidopsi
26	101.5	5.5	1016	3	Q9H745	Q9H745 schizosach
27	99	5.4	319	5	Q9Z632	Q9Z632 caenorhabdi
28	98	5.3	1236	5	Q44875	Q44875 caenorhabdi
29	97	5.3	1236	5	Q44875	Q44875 caenorhabdi
30	96	5.2	1280	2	Q9ZAJ5	Q9ZAJ5 cystridium
31	96	5.2	1862	10	Q9S8M5	Q9S8M5 arabidopsi
32	94.5	5.1	469	14	Q9W852	Q9W852 physalis se
33	94.5	5.1	469	14	Q9W852	Q9W852 physalis se
34	94.5	5.1	787	1	Q9H925	Q9H925 pyrococcus
35	94.5	5.1	787	1	Q9H925	Q9H925 pyrococcus
36	94.5	5.1	787	1	Q9H925	Q9H925 pyrococcus
37	94	5.1	4442	14	Q9ZUK3	Q9ZUK3 turkey harp
38	94	5.1	4442	14	Q9ZUK3	Q9ZUK3 turkey harp
39	93.5	5.1	525	3	Q9Z830	Q9Z830 schizosach
40	93	5.0	735	2	Q9XMK6	Q9XMK6 bacteroides
41	93	5.0	749	5	Q9Z602	Q9Z602 plasmodium
42	93	5.0	897	10	Q91H00	Q91H00 arabidopsi
43	92.5	5.0	2185	14	Q9E7C3	Q9E7C3 human coxa
44	92.5	5.0	2185	14	Q9E7C3	Q9E7C3 human coxa
45	92	5.0	392	3	Q94240	Q94240 schizosach

#### ALIGNMENTS

RESULT 1  
Q9FZNB 1 PPIIMINAVY: PPT: 369 AA.

ID Q9FZNB: PPIIMINAVY: PPT: 369 AA.  
AC Q9FZNB: PPIIMINAVY: PPT: 369 AA.  
DT 01-MAR-2001 (TREMBLrel. 16, Created)  
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
DE CAFFEINE SYNTHASE.  
CN TCSI.  
OS Camellia sinensis (Tea).  
OR Enayyati, Viridiplantae, Eriythyta, Tracheophyta, Spermatophyta,  
OC Malvophyta, eudicotyledons, core eudicotyledons, Asterales, Ericales)  
OC Theaceae; Camellia.  
OX NCBI\_TaxID=4442;  
PN [1]  
PP SEQUENCE FROM N.A.  
RA MEDLINE=20437335; PubMed=10984041;  
RA Kato M., Mizuno K., Crozier A., Fujimura T., Ashihara H.:  
RT "Caffeine synthase gene from tea leaves."  
RU Nature 406:956-957(2000).  
DR EMBL: AB031280; EMBL227811; -; F01065874/5576 r0664;  
SQ SEQUENCE: 369 AA: 4127 MW: 41277 MW: 41277 MW: 41277

Query Match 100.0% Score 1847: OR 19: Length 369;  
Best Local Similarity 100.0%: Pred. No. 1, 1e-144;  
Matches 355: Conservative 0: Mismatches 0: Indels 0: Gaps 0:  
Q7 1 FNNRPPSSSYAANSSTFQ2VASMADPALENVETLPSMDPILQAINADLCAAGPNTF 60  
|||||  
Q7 14 FNNRPPSSSYAANSSTFQ2VASMADPALENVETLPSMDPILQAINADLCAAGPNTF 73  
|||||  
Q7 61 AVTSTFEMKPKFRETNETFLQVYNTLFESNGRIHFKLSSVVGSRKEVGVVM 120  
|||||  
Q7 74 AVTSTFEMKPKFRETNETFLQVYNTLFESNGRIHFKLSSVVGSRKEVGVVM 133  
|||||  
Q7 121 GVPSPSPDRIEKNSEDEVSSYSVMEIPANKSLSPEDALNKGKLYSKSPVVE 180  
|||||  
Q7 134 GVPSPSPDRIEKNSEDEVSSYSVMEIPANKSLSPEDALNKGKLYSKSPVVE 193  
|||||  
Q7 181 AVTSTFEMKPKFRETNETFLQVYNTLFESNGRIHFKLSSVVGSRKEVGVVM 240  
|||||





HA Tabata S.:  
 RT "Structural analysis of Arabidopsis thaliana chromosome 5. VI  
 RT Sequence features of the regions of 1,367,183 bp covered by 14  
 RT physically assayed P1 and PAC clones."  
 DR DNA Res. 5:293-316(1998).  
 DR EMBL: A013449; F010919.1; -  
 SO SEQUENCE 354 AA: 40033 MW: 9519636DCACAD329 CRC64;

Query Match 34.58; Score 637.5; DB 10; Length 354;  
 Best Local Similarity 39.78; Pred. No. 9.7; 45;  
 Matches 140; Conservative 64; Mismatches 131; Indels 25; Gaps 9;

QY 2 MNRGSSVAQNSSTFGQVASMAGPAP-ENAVETI PSRPHH QAINADLGAAGPNT 60  
 DB 1 MSGGQNSSTNSI LKRVLSKAPVIVNTGMMINLNP NTKVADLGCATGENT 59  
 QY 61 AVISTIKMMKRCRELNCQLEQVYVNDLPNDENTLFGKLSSEVIGKCEVPC 117  
 DB 60 LTMALVNTLVNLCQCNKRPETDCLNLPNDENTLFGKVPFNKRVKSKPI- 115  
 QY 118 YVMGVPSSFRHLEFPNSLHVSSSVHMLTAPKGLTREGALNKKGIYSKSPV 177  
 DB 116 FVSGVPSSFSPI FPKSLFHVSSSVHMLSKVPKGLK-----NSSVYITTSPPN 169  
 QY 178 VREAVLSQFHDFTPLNASSQFVPCVWLLPGHQ-CDSIDMQSCFTWELLAMATA 236  
 DB 170 AYKAVLNQFQDSKSEFLKMSSEKVNKGPVLLTFIGKTLIDPLHDCHEWLLSTSLP 229  
 QY 237 ELVSQGLIDELKLTENIPSYFASLEVKDIVERDGSFTIDHLE-GFDL----DSVHM 289  
 DB 230 DLVEGLVASAKVSDFNIPEDPSKEVMDIENEGSEINDLEIGHFELGSLNHEDEYA 289  
 QY 290 QFNKRWPKCEKFTVVAFAFTEPIISNGFPEIMKLYDKFTLVVSCITAKLPTTPI 349  
 DB 290 LHSOISAGQKFAKCTRAVSHMVALFQVIMLELKKFAHV--SGHASTKNTKTVH 347  
 QY 350 VLKRI 354  
 DB 348 VVSLI 352  
 RESULT 5  
 ID 023234 PRELIMINARY; PRT: 371 AA.  
 AC 023234;  
 DT 01-JAN-1998 (TEMBLrel. 05, Created)  
 DT 01-MAY-1999 (TEMBLrel. 10, Last sequence update)  
 DT 01-OCT-2000 (TEMBLrel. 15, Last annotation update)  
 DB HYPOTHETICAL 42.0 KDA PROTEIN.  
 GN C7A10.890 OR A7436470  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
 OC Magnoliophyta; Euphorbiales; Ericaceae; Vaccinioideae; Ericaceae;  
 OC Brassicales; Brassicaceae; Arabidopsids.  
 OC NCBI\_TaxID=3702;  
 RN 11  
 RP SEQUENCE FROM N.A.  
 RA Bryan M. Terryn N. Vos P. Heijnen L. Meuwis H.W. Scholten C.  
 RA Chabalaz N.;  
 PI Submitted (Dec-1998) to the EMBL/GenBank/DDBI databases.  
 RN 12  
 RP SEQUENCE FROM N.A.  
 RA EU Arabidopsis sequencing project;  
 RL Submitted (MAR 2000) to the EMBL/GenBank/DDBI databases.  
 DR EMBL: Z99708; CAB6845.1; -  
 DR EMBL: A1161589; CAB8013.1; -  
 KW Hypothetical protein.  
 SO SEQUENCE 371 AA: 41984 MW: F5PR341ERSGRCA90 CR064.

Query Match 34.18; Score 630; DB 10; Length 371;  
 Best Local Similarity 37.38; Pred. No. 4.3e-44;

Matches 135; Conservative 84; Mismatches 131; Indels 12; Gaps 5;

QY 1 FNNQSSVAQNSSTFGQVASMAGPAP-ENAVETI PSRPHH QAINADLGAAGPNT 60  
 DB 1 YMGQDKISAPNSIQRKASTAKHTITLQQLY-KEIPFSQIMLQSSSPNT 69  
 QY 61 AVISTIKMMKRCRELNCQLEQVYVNDLPNDENTLFGKLSSEVIGKCEVPC 117  
 DB 70 STTDFKTVGVVHHPETPIQPLEPSTPLNPPNDENTLFGKVPFNKRVKSKPI- 129  
 QY 112 GFVNVVGVSSFRHLEFPNSLHVSSSVHMLTAPKGLTREGALNKKGIYS 171  
 DB 130 GPSV--FLAAYPGSFGVRLPEPNTIHVVASSTHMLSNVPLATYFGKSLNKQVSLC 187  
 QY 172 KTSFTVFAVLSQFHDFTPLNASSQFVPCVWLLPGHQ-CDSIDMQSCFTWELL 231  
 DB 188 SLSEAVSKAVSQKEDPSTFLPRSKEVASGPVLLTIGRPPHVPQNSFEWELL 247  
 QY 232 AMALVELSQGLIDELKLTENIPSYFASLEVKDIVERDGSFTIDHLE-GFDL----DSVHM 289  
 DB 248 SRSLADLVAGQFTFRKIDSYDHPHFAPSADFTGCVKRGSPFLRDLMLKDKKNT 307  
 QY 292 NKKWVPCEKFTKVVAFTEPIISNGFPEIMKLYDKFTLVVSCITAKLPTTPI 351  
 DB 308 FQTSYCAVAKIVFAVSHMVALFQVIMLELKKFAHV--SGHASTKNTKTVH 367  
 QY 352 SK 353  
 DB 368 RK 369  
 RESULT 6  
 ID 09FY29 PRELIMINARY; PRT: 364 AA.  
 AC 09FY29;  
 DT 01-MAR-2001 (TEMBLrel. 16, Created)  
 DT 01-MAR-2001 (TEMBLrel. 16, Last sequence update)  
 DT 01-MAR-2001 (TEMBLrel. 16, Last annotation update)  
 DB SAM-BENZOIC ACID CARBOXYL METHYLTRANSFERASE.  
 GN BANT.  
 OS Antirrhinum majus (Garden snapdragon).  
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
 OC Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I;  
 OC Lamiales; Scrophulariaceae; Antirrhinum.  
 OC NCBI\_TaxID=4151;  
 RN 111  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=20312758; PubMed=10852939;  
 RA Dzhirava N., Morfitt I.M., Mann C.T., Gredeslein N., Kolesova N.,  
 RA Kish C.M., Bonham C., Wood K.;  
 PT "Developmental regulation of methyl benzoate biosynthesis and emission  
 RT in snapdragon flowers."  
 RL Plant Cell 12:949-961(2000).  
 LE EMBL: A136452; AA:0584.1; -  
 KW Transferase, Methyltransferase.  
 SO SEQUENCE 364 AA: 41010 MW: C10C8E86A591419 CR064;

Query Match 33.68; Score 621.5; DB 10; Length 364;  
 Best Local Similarity 38.18; Pred. No. 2.1e-43;  
 Matches 140; Conservative 64; Mismatches 128; Indels 35; Gaps 8;

QY 5 GERSSVAQNSSTFGQVASMAGPAP-ENAVETI PSRPHH QAINADLGAAGPNT 58  
 DB 14 GTSSTANNSLQKVMKSLHVLPTIKLIIQHVHPPKPKKM-----LMSGSSGPN 68  
 QY 59 TFAVSTIKMMKRCRELNCQLEQVYVNDLPNDENTLFGKLSSEVIGKCEVPC 117  
 DB 69 ALVMSCLINIEELYTEKINELLEFEVHLNLPNDENRDLKLSH RGN - - - C 121  
 QY 118 YVMGVPSSFRHLEFPNSLHVSSSVHMLTAPKGLTREGALNKKGIYSKSPV 177  
 DB 122 FVGLPSFSVGLPKSLFAVSYSHMLSOVPLEDN-----NNQNTYMAIESPPE 176





```

QY 5 GPCSSAONSSFTQGVASMAQALFNAVEITLPSDFHIALNALALQAAQNPFAVIS 64
DB 10 GRPNSTYFHSKRYGAI VIAAKERINAEALSTKLDIDFISNLVINDFGSSQNPFLF 69
QY 65 TIKRMRKRC-PRI NGTLELVYNDLEGNDFNTL PKRI SSEVITONKFFVPCVPCV 123
DB 70 TLIAVANIKKRSKSNLEGIHFQVFNDSNNDFNTL FKLTPARL-----YFASVCP 121
QY 124 GSTRHGLFPNSILVHSSSVHMLTQAPKQLTSPELALNCGYVYKSTSPVAVL 123
DB 122 GSFEGRVIFPNSLHVSSSVHMLTQAPKQLTSPELALNCGYVYKSTSPVAVL 123
QY 184 SOFHEPTEMLNARSGEVVNGMVLIRGROSDPSDMSQFTWE-----LMAAIADL 238
DB 180 GYKLLVGSFLTARAGELVSGILLILLSGR-----PFGVMEIYEGMIDPISLSMEI 235
QY 239 VSGTITPFTFENPSPASTFVYVIVPFGQSTFD-----HIEG-EDLSVEKCE 291
DB 236 ANGLIDOCKITPFPKLPVAPNVETNLIETENKCHLEAFERKISHAGCEYELDP----- 290
QY 292 NDKWVREKRTKVVAFTEFLISNPEFIMETRYKPHIVASDIPATIPK 343
DB 291 -----EYLSAFKVTGVGSVASLFGVGMERTF-----LVKERTQEMLPQ 331

```

## RESULT 12

Q9FYC6 PRELIMINARY: PRT: 348 AA.

```

ID 09FYC6:
AC 09FYC6:
DT 01-MAR-2001 (TREMUR, 16, Created)
DT 01-MAR-2001 (TREMUR, 16, last sequence update)
DT 01-MAR-2001 (TREMUR, 16, last annotation update)
DE PROTEININASE ATP-LIKE PROTEIN
GN F28D10.30
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicotids; Rosidae; eucosids II;
OC Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN 111
RP SEQUENCE FROM N.A.
RA Delseny M., Berger C., Cooke R., Grellet F., Laurie M., Mewes H.W.,
RA Roud S., Iwano K., Mayer K.F.X., Queller F., Salanoubat M.;
RL Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.
RN 121
RP SEQUENCE FROM N.A.
RA FU Arabidopsis sequencing project;
RL Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases
DR EMBL; AL391254; CAC01534.1; -.
KW Kinase.
SQ SEQUENCE 348 AA: 48940 MW: 41600FFPEVETAT4 TCT4.

```

Query Match 26.78; Score 485.5; DB 10; Length 348,  
Host Local Similarity 44.94; Pred No. 3; Size 32;  
Matches 128; Conservative 65; Mismatches 129; Indels 45; Gaps 13;

```

QY 5 GPCSSAONSSFTQGVASMAQALFNAVEITLPSDFHIALNALALQAAQNPFAVIS 64
DB 10 GRPNSTYFHSKRYGAI VIAAKERINAEALSTKLDIDFISNLVINDFGSSQNPFLF 69
QY 65 TIKRMRKRC-PRI NGTLELVYNDLEGNDFNTL PKRI SSEVITONKFFVPCVPCV 123
DB 70 TLIAVANIKKRSKSNLEGIHFQVFNDSNNDFNTL FKLTPARL-----YFASVCP 121
QY 124 GSTRHGLFPNSILVHSSSVHMLTQAPKQLTSPELALNCGYVYKSTSPVAVL 123
DB 122 GSFEGRVIFPNSLHVSSSVHMLTQAPKQLTSPELALNCGYVYKSTSPVAVL 123
QY 184 SOFHEPTEMLNARSGEVVNGMVLIRGROSDPSDMSQFTWE-----LMAAIADL 238
DB 180 GYKLLVGSFLTARAGELVSGILLILLSGR-----PFGVMEIYEGMIDPISLSMEI 235
QY 239 VSGTITPFTFENPSPASTFVYVIVPFGQSTFD-----HIEG-EDLSVEKCE 291
DB 236 ANGLIDOCKITPFPKLPVAPNVETNLIETENKCHLEAFERKISHAGCEYELDP----- 290
QY 292 NDKWVREKRTKVVAFTEFLISNPEFIMETRYKPHIVASDIPATIPK 343
DB 291 -----EYLSAFKVTGVGSVASLFGVGMERTF-----LVKERTQEMLPQ 331

```

```

QY 235 IALVSGCLIDELKIDTNPVPSYFASLEVKDVIKDSFTIDHTEGPDLSVMKCHIK 294
DB 232 LNDIAGGLIDELKIDTNPVPSYFASLEVKDVIKDSFTIDHTEGPDLSVMKCHIK 285
QY 295 WVGFEFTFY-VPAFPEITSMVPGPEIMDKLYCFETHVASEVAKLPKTS-----I 347
DB 286 IPIIDPEITVPSFVTVGGIVASPGQGVMEKTP-----VVKIKTEPMI.POLANKPCMOY 341
QY 342 LVVSTI 354
DB 342 LVVSTI 348

```

## RESULT 13

Q9LS10 PRELIMINARY: PRT: 459 AA.

```

ID 09LS10:
AC 09LS10:
DT 01-OCT-2000 (TREMUR, 15, Created)
DT 01-OCT-2000 (TREMUR, 15, last sequence update)
DT 01-MAR-2001 (TREMUR, 16, last annotation update)
DE S-ADENOSYL-L-METHIONINE:SALICYLIC ACID C-ABEAXYL METHYLTRANSFERASE-
DE LIKE
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicotids; Rosidae; eucosids II;
OC Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN 111
RP SEQUENCE FROM N.A.
RA STRAIN=COLUMBIA;
RA Kaneko T., Katoh T., Asamizu H., Sato S., Nakamura Y., Kotani H.,
RA Tabata S.;
RL Structural analysis of Arabidopsis thaliana chromosome 5. XI.*;
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB028060; BA097544.1; -.
SQ SEQUENCE 553 AA: 40944 MW: 456527A80950D34 CROG4;

```

Query Match 26.18; Score 482.5; DB 10; Length 459;  
Host Local Similarity 33.78; Pred. No. 6; Size 32;  
Matches 115; Conservative 65; Mismatches 124; Indels 47; Gaps 11;

```

QY 2 MNRGESSAONSSFTQGVASMAQALFNAVEITLPSDFHIALNALALQAAQNPFAVIS 56
DB 10 MNRGESSAONSSFTQGVASMAQALFNAVEITLPSDFHIALNALALQAAQNPFAVIS 56
QY 57 PNFV-----AVSTIKRMRKRC-PRI NGTLELVYNDLEGNDFNTL PKRI SSEVITONKFFVPCVPCV 110
DB 67 PNFV-----AVSTIKRMRKRC-PRI NGTLELVYNDLEGNDFNTL PKRI SSEVITONKFFVPCVPCV 110
QY 112 KLELVYVMAVISEHDLFPNSLHVSSSVHMLTQAPKQLTSPELALNCGYVYKSTSPVAVL 170
DB 121 -PEPEFVSQGVSGVGLPNSLHVSSSVHMLTQAPKQLTSPELALNCGYVYKSTSPVAVL 176
QY 171 SKSTIVVFPVAT SQPHDPTM; NARSGVYVNGMVLIRGROSDPSDMSQFTWE-----LMAAIADL 230
DB 179 NNLLEEVTK-AKVQPEIKMEIFLAPAEELVPGGLMVI---GEGL-PCGVSLYETWOG 234
QY 231 LAR-----ALVLSQGLIDELKIDTNPVPSYFASLEVKDVIKDSFTIDHTEGPDLSVMKCHIK 294
DB 234 YVMDITGGLLMARSGVGLPNSLHVSSSVHMLTQAPKQLTSPELALNCGYVYKSTSPVAVL 286
QY 296 SVPMENDRWVREKRTKVVAFTEFLISNPEFIMETRYKPHIVASDIPATIPK 343
DB 289 ITSHPLEGRPLNDFTITFPAFLITLIEKHGQDVVELF 329

```

## RESULT 14

Q9FYC6 PRELIMINARY: PRT: 374 AA.

```

ID 09FYC6:
AC 09FYC6:
DT 01-MAR-2001 (TREMUR, 16, Created)
DT 01-MAR-2001 (TREMUR, 16, last sequence update)

```



Sat Sep 1 11:45:37 2001

us-09-577-657a-1.rpt

---

Page 8